

2017 데이터 산업 현황조사





승인번호
제 127004 호

발간등록번호

11-171000-000387-10

2017 데이터 산업 현황조사



과학기술정보통신부 Kdata 한국데이터진흥원



2017 데이터 산업 현황 조사



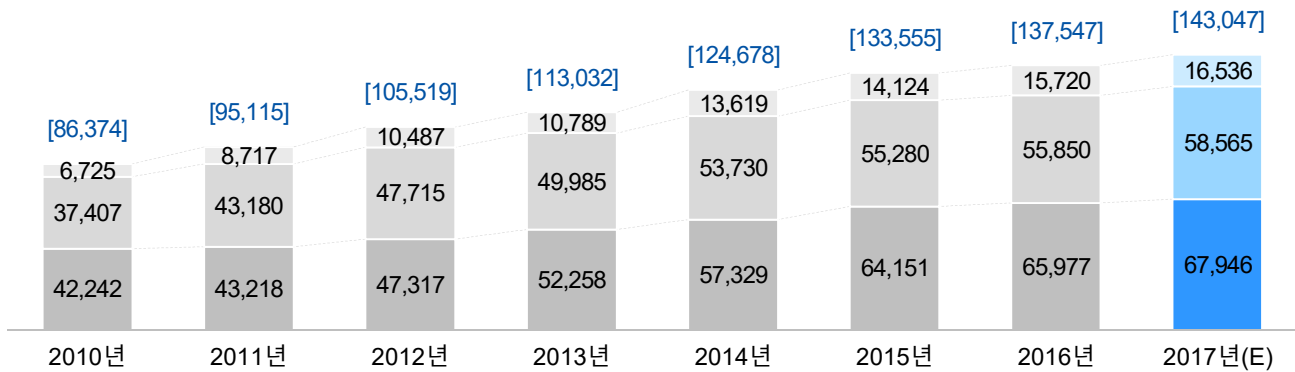
과학기술정보통신부 Kdata 한국데이터진흥원

그림으로 보는 데이터산업 동향

1. 데이터산업 시장 규모

'10~'17 CAGR 7.5%

■ 데이터솔루션 ■ 데이터구축/컨설팅 ■ 데이터서비스 [전체] (단위: 억 원)

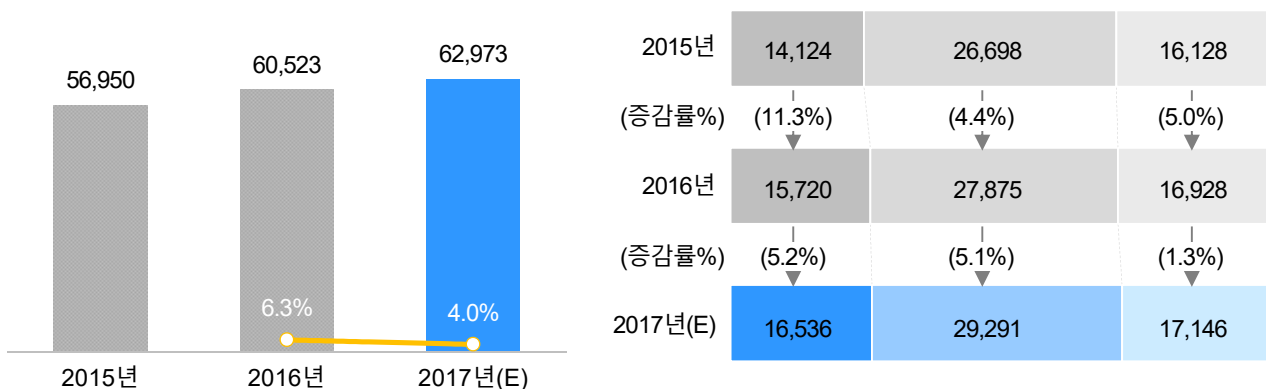


※ 2016년은 확정치이므로 2016년 조사 결과의 2016년(E)(예상치) 수치와 서로 상이함

※ 데이터산업 시장은 전통적인 DB산업을 아우르는 기존 조사 범위를 적용한 것으로, 데이터를 매개로하는 광고 매출, DB시스템 구축 용역 매출 등 DB 관련 간접 매출까지 포함하여 산출한 광의의 데이터산업 시장 규모임

2. 데이터산업 직접매출 시장 규모

■ 데이터산업 직접매출 (단위: 억 원) ■ 데이터솔루션 ■ 데이터구축/컨설팅 ■ 데이터서비스 (단위: 억 원)
○ 증감률

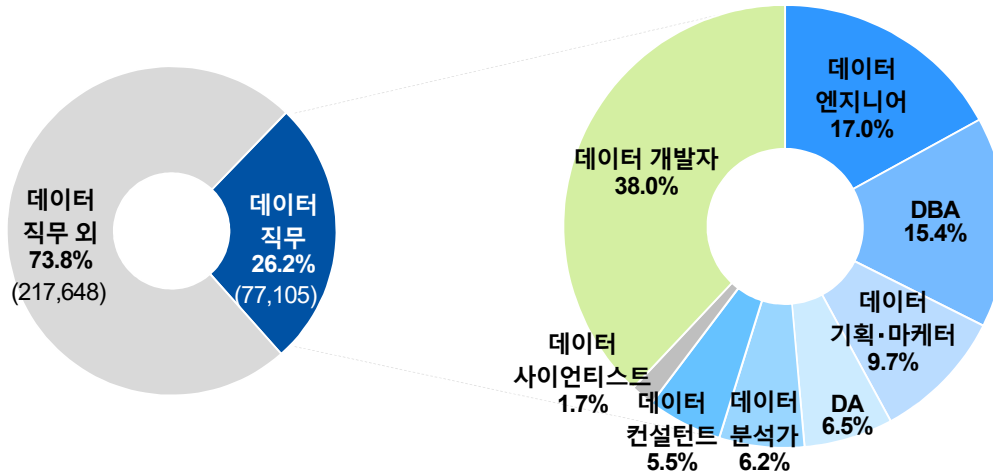


※ 데이터산업 직접매출 시장 규모는 데이터와 직접 연관이 되는 매출을 기준으로 조사한 협의의 데이터산업 시장 규모임

그림으로 보는 데이터산업 동향

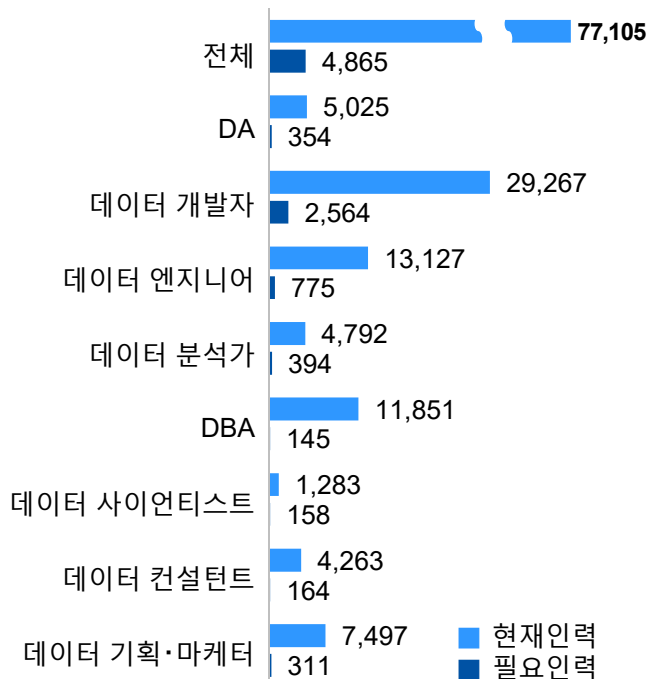
3. 데이터산업 인력 현황

2017년 데이터산업 인력: 총 294,753명

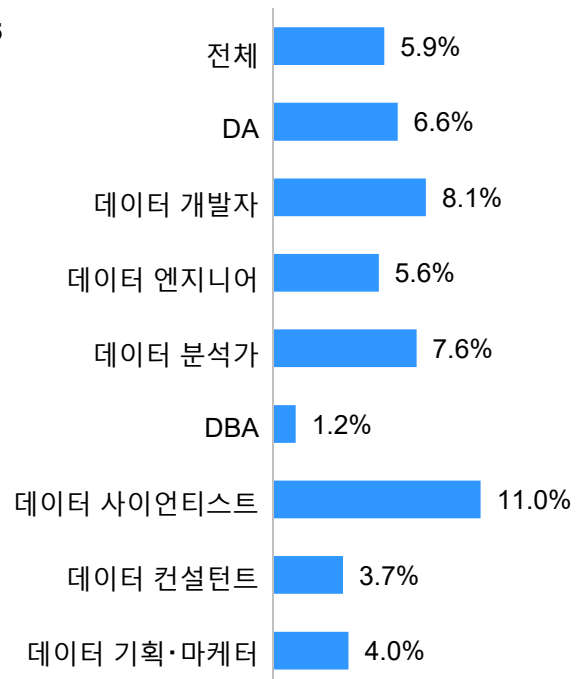


4. 데이터산업의 데이터직무 인력 현황 및 부족률

※ 2017년 데이터산업의 데이터직무 인력: 총 77,105명



※ 2017년 데이터산업의 데이터직무별 인력 부족률



2017년 데이터산업 현황 조사 결과 요약

■ 조사 개요

- 조사목적 : 데이터산업 환경 변화 파악 및 체계적인 데이터산업 육성 지원 정책 수립을 위한 기초 통계 제공
- 조사대상 : 전국에 소재한 데이터산업 영위 사업체 7,133개
(표본수 1,000개, 조사응답기업 1,004개, 회수율 100.4%)
- 조사내용 : 데이터 관련 사업 매출, 인력 현황 및 수요 등
- 조사방법 : 구조화된 설문지를 이용하여 방문면접조사를 기본으로 실시
- 조사기간 : 2017. 10. ~ 2017. 12.

■ 주요 조사 결과

[1] 조사 사업체 현황

- 2017년 국내 데이터산업 사업체 수는 7,133개이며, 이 중 종사자 수 1~10인 미만 규모의 사업체가 73.5%로 가장 큰 비중을 차지하는 것으로 나타났다.

[표1] 데이터산업 사업체수 현황

구분	1-9인	10-49인	50-99인	100- 299인	300인 이상	합계	비중
데이터솔루션	1,183	289	73	45	18	1,608	22.6%
데이터구축/컨설팅	1,827	538	99	93	41	2,598	36.4%
데이터서비스	2,232	466	98	90	41	2,927	41.0%
합계	5,242	1,293	270	228	100	7,133	100.0%
비중	73.5%	18.1%	3.8%	3.2%	1.4%	100.0%	

(단위: 억 원)

[2] 데이터산업 시장 규모¹⁾

가. 전체 시장 규모

- 국내 데이터산업 시장 규모는 2016년 13조 7,547억 원이며, 2017년은 14조 3,047억 원으로 전년 대비 4.0% 성장할 것으로 나타났다. 2010년 이후 연평균증가율은 7.5%로 나타났다.

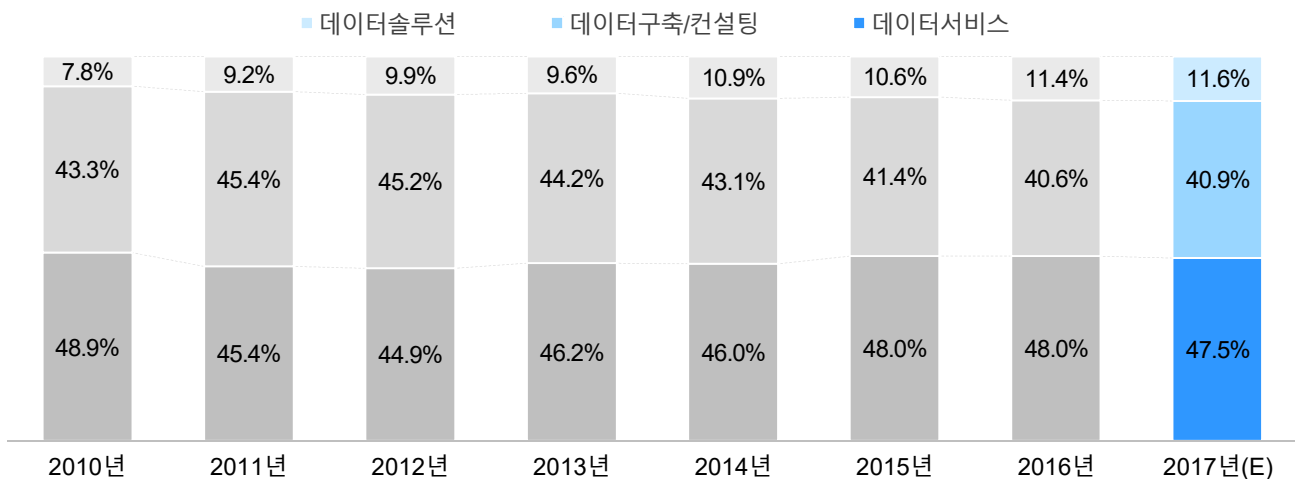
[표2] 데이터산업 시장 규모

구분	2010년	2011년	2012년	2013년	2014년	2015년	2016년	2017년 (E)	증감률 '16~'17	CAGR '10~'17
데이터솔루션	6,725	8,717	10,487	10,789	13,619	14,124	15,720	16,536	5.2%	13.7%
데이터 구축/컨설팅	37,407	43,180	47,715	49,985	53,730	55,280	55,850	58,565	4.9%	6.6%
데이터서비스	42,242	43,218	47,317	52,258	57,329	64,151	65,977	67,946	3.0%	7.0%
전체	86,374	95,115	105,519	113,032	124,678	133,555	137,547	143,047	4.0%	7.5%

※ 2016년은 확정치이므로 2016년 조사 결과의 2016년(E)(예상치) 수치와 서로 상이함

(단위: 억 원)

[그림1] 데이터산업 부문별 시장 규모 비중



1) 데이터산업 시장은 전통적인 DB산업을 아우르는 기존 조사 범위를 적용한 것으로, 데이터를 매개로하는 광고 매출, DB시스템 구축 용역 매출 등 DB 관련 간접 매출까지 포함하여 산출한 광의의 데이터산업 시장 규모임

나. 부문별 시장 규모

1) 데이터솔루션

- 데이터솔루션 시장 규모는 2016년 1조 5,720억 원이며, 2017년에는 1조 6,536억 원으로 전년 대비 5.2% 성장할 것으로 나타났다.

[표3] 데이터솔루션 중분류별 시장 규모

구분	2010년	2011년	2012년	2013년	2014년	2015년	2016년	2017년 (E)	증감률 '16~'17	CAGR '10~'17
데이터 수집	448	570	617	622	1,076	1,115	1,345	1,382	2.8%	17.5%
데이터 설계	129	136	151	166	202	207	225	269	19.6%	11.1%
DBMS	3,169	4,272	5,301	5,488	5,502	5,727	6,148	6,192	0.7%	10.0%
데이터 관리	1,237	1,516	2,075	2,489	3,446	3,574	4,022	4,417	9.8%	19.9%
데이터 품질	686	949	942	855	883	918	1,120	1,127	0.6%	7.3%
데이터 분석	1,056	1,274	1,401	1,169	1,121	1,157	1,249	1,328	6.3%	3.3%
데이터 플랫폼	-	-	-	-	1,389	1,426	1,611	1,821	13.0%	9.4% ²⁾
데이터솔루션 전체	6,725	8,717	10,487	10,789	13,619	14,124	15,720	16,536	5.2%	13.7%

(단위: 억 원)

2) 데이터구축 및 컨설팅

- 데이터구축 시장은 2016년 5조 4,571억 원이며, 2017년은 5조 6,893억 원으로 전년 대비 4.3% 성장할 것으로 나타났다. 데이터컨설팅 시장은 2016년 1,279억 원이며, 2017년에는 1,672억 원으로 2016년 대비 30.7% 성장할 것으로 조사되었다.

[표4] 데이터구축 및 컨설팅 시장 규모

구분	2010년	2011년	2012년	2013년	2014년	2015년	2016년	2017년 (E)	증감률 '16~'17	CAGR '10~'17
데이터 구축	36,610	42,374	46,865	49,021	52,673	54,142	54,571	56,893	4.3%	6.5%
데이터 컨설팅	797	806	850	964	1,057	1,138	1,279	1,672	30.7%	11.2%
데이터 구축/컨설팅 전체	37,407	43,180	47,715	49,985	53,730	55,280	55,850	58,565	4.9%	6.6%

(단위: 억 원)

2) 2014년~2017년 CAGR

3) 데이터서비스

- 데이터서비스 시장은 2016년 6조 5,977억 원이며, 2017년에는 6조 7,946억 원으로 전년 대비 3.0% 증가할 것으로 조사되었다. 데이터서비스 세부분야별 증감률은 데이터거래가 4.4%, 데이터 분석제공이 3.8%, 정보제공이 2.9% 순으로 나타났다.

[표5] 데이터서비스 중분류별 시장 규모

구분	2013년		2014년		2015년		2016년		2017년(E)		증감률 '16~'17	CAGR '13~'17
	규모	비중	규모	비중	규모	비중	규모	비중	규모	비중		
데이터거래	1,708	3.3%	2,019	3.5%	2,379	3.7%	2,594	3.9%	2,709	4.0%	4.4%	12.2%
정보 제공	48,284	92.4%	51,985	90.7%	58,171	90.7%	59,854	90.7%	61,575	90.6%	2.9%	6.3%
데이터 분석 제공	2,266	4.3%	3,325	5.8%	3,601	5.6%	3,529	5.4%	3,662	5.4%	3.8%	12.7%
데이터서비스 전체	52,258	100.0%	57,329	100.0%	64,151	100.0%	65,977	100.0%	67,946	100.0%	3.0%	6.8%

(단위: 억 원)

[3] 데이터산업 직접매출 시장 규모

가. 전체 시장 규모

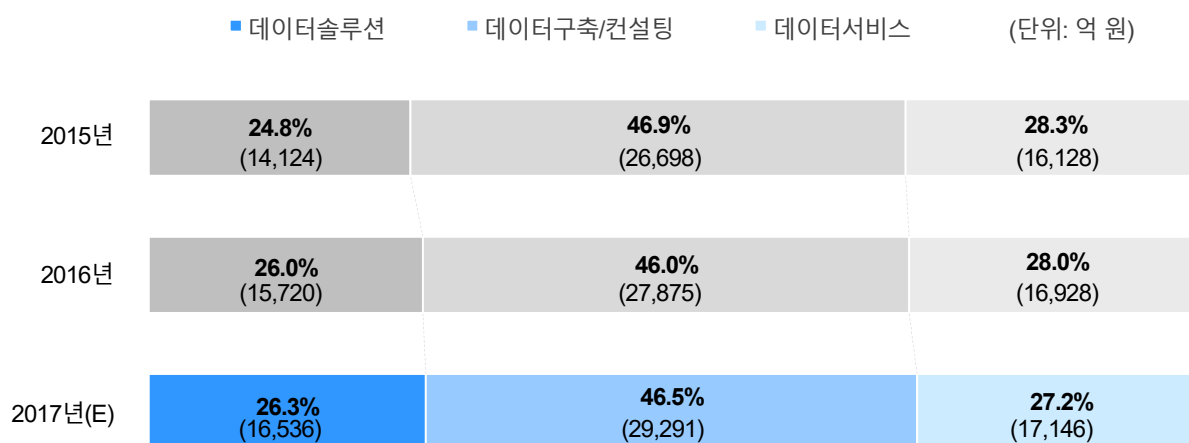
- 국내 데이터산업 직접매출 시장 규모는 2016년 6조 523억 원이며, 2017년에는 6조 2,973억 원으로 전년 대비 4.0% 성장할 것으로 조사되었다.

[표6] 데이터산업 부문별 직접매출 시장 규모

구분	2015년		2016년		2017년(E)		증감률 '16~'17
	시장 규모	비중	시장 규모	비중	시장 규모	비중	
데이터솔루션	14,124	24.8%	15,720	26.0%	16,536	26.3%	5.2%
데이터구축/컨설팅	26,698	46.9%	27,875	46.0%	29,291	46.5%	5.1%
데이터서비스	16,128	28.3%	16,928	28.0%	17,146	27.2%	1.3%
전체	56,950	100.0%	60,523	100.0%	62,973	100.0%	4.0%

(단위: 억 원)

[그림2] 데이터산업 부문별 직접매출 시장 규모 비중



나. 부문별 시장 규모

○ 국내 데이터산업 세부영역별 직접매출 시장 규모는 다음과 같다.

[표7] 데이터산업 세부영역별 직접매출 시장 규모

구분		2015년	2016년	2017년(E)	증감률 '16~'17
데이터 솔루션	데이터 수집	1,115	1,345	1,382	2.8%
	데이터 설계	207	225	269	19.6%
	DBMS	5,727	6,148	6,192	0.7%
	데이터 관리	3,574	4,022	4,417	9.8%
	데이터 품질	918	1,120	1,127	0.6%
	데이터 분석	1,157	1,249	1,328	6.3%
	데이터 플랫폼	1,426	1,611	1,821	13.0%
데이터솔루션 전체		14,124	15,720	16,536	5.2%
데이터 구축/컨설팅	데이터구축	25,560	26,596	27,619	3.8%
	데이터컨설팅	1,138	1,279	1,672	30.7%
데이터구축/컨설팅 전체		26,698	27,875	29,291	5.1%
데이터 서비스	데이터 거래	2,350	2,354	2,409	2.3%
	정보 제공	10,726	11,523	11,632	0.9%
	데이터 분석 제공	3,052	3,051	3,105	1.8%
데이터서비스 전체		16,128	16,928	17,146	1.3%
전체		56,950	60,523	62,973	4.0%

(단위: 억 원)

(4) 데이터직무 인력 현황 및 수요

가. 데이터직무 전체

1) 데이터직무 인력 현황

- 일반산업을 포함해 전 산업의 데이터직무 인력은 2017년 총 10만 9,320명으로 전년 대비 6.8% 증가하였다.³⁾

[표8] 전 산업의 데이터직무 인력 현황 추이

구분	2015년		2016년		2017년		증감률 '16~'17
	규모(명)	비중(%)	규모(명)	비중(%)	규모(명)	비중(%)	
데이터산업	70,338	70.0%	73,256	71.6%	77,105	70.5%	5.3%
일반산업	30,102	30.0%	29,119	28.4%	32,215	29.5%	10.6%
전 산업	100,440	100.0%	102,375	100.0%	109,320	100.0%	6.8%

(단위: 명, %)

- 데이터직무별로는 데이터 개발자가 2017년 4만 1,254명으로 가장 큰 비중을 차지하며, 이어서 DBA가 1만 7,863명, 데이터 엔지니어가 1만 6,634명 순으로 나타났다.

[표9] 2017년 전 산업의 데이터직무별 인력 현황

구분	데이터산업					일반산업	전체 (데이터+일반)
	데이터 솔루션	데이터 구축	데이터 컨설팅	데이터 서비스	소계		
DA	618	3,392	526	489	5,025	5,046	10,071
데이터 개발자	4,429	13,702	2,188	8,948	29,267	11,987	41,254
데이터 엔지니어	1,468	6,529	696	4,434	13,127	3,507	16,634
데이터 분석가	629	1,638	335	2,190	4,792	3,606	8,398
DBA	702	3,732	388	7,029	11,851	6,012	17,863
데이터 사이언티스트	218	410	173	482	1,283	520	1,803
데이터 컨설턴트	1,530	1,376	1,231	126	4,263	741	5,004
데이터 기획·마케터	697	914	286	5,600	7,497	796	8,293
전체	10,291	31,693	5,823	29,298	77,105	32,215	109,320

(단위: 명)

3) 일반산업 모집단 및 조사표본 수는 다음과 같음

구분	모집단 기업 수 (종사자 100명 이상)	2016년		2017년	
		조사표본 수	비중	조사표본 수	비중
공공	1,157	202	23.1%	164	18.7%
금융	399	43	4.9%	44	5.0%
유통/서비스	1,916	146	16.7%	186	21.3%
제조	4,195	357	40.9%	353	40.3%
의료	417	64	7.3%	74	8.5%
통신/미디어	1,325	62	7.1%	54	6.2%
일반산업 전체	9,409	874	100.0%	875	100.0%

※ 2017년 일반산업 결과는 '2017년 빅데이터 시장 현황조사' 데이터를 활용한 결과임

- 2016년 대비 가장 높은 증가율을 나타낸 데이터직무는 데이터 분석가로서 2017년은 전년 대비 14.4% 증가한 8,398명인 것으로 조사되었다.

[표10] 2016년-2017년 전 산업의 데이터직무별 인력 현황 비교

구분	2016년		2017년		증감률 (‘16~’17)
	규모(명)	비중(%)	규모(명)	비중(%)	
DA	9,267	9.1%	10,071	9.2%	8.7%
데이터 개발자	38,948	38.0%	41,254	37.7%	5.9%
데이터 엔지니어	15,670	15.3%	16,634	15.2%	6.2%
데이터 분석가	7,339	7.2%	8,398	7.7%	14.4%
DBA	17,116	16.7%	17,863	16.3%	4.4%
데이터 사이언티스트	1,662	1.6%	1,803	1.7%	8.5%
데이터 컨설턴트	4,513	4.4%	5,004	4.6%	10.9%
데이터 기획·마케터	7,860	7.7%	8,293	7.6%	5.5%
전체	102,375	100.0%	109,320	100.0%	6.8%

(단위: 명, %)

2) 데이터직무 인력 수요

- 향후 3년 내 추가로 필요한 데이터직무 인력은 총 1만 3,337명이며, 이 중 데이터 개발자 (37.2%)의 인력 부족률이 가장 높았고, 다음으로 DA(13.4%)와 데이터 엔지니어(13.0%) 등의 순으로 부족률이 높게 나타났다.

[표11] 2017년 전 산업 내 데이터직무별 필요인력 및 인력 부족률

구분	데이터산업 필요인력	일반산업 필요인력	전 산업 필요인력 (데이터산업+일반산업)		인력부족률
			규모(명)	비중(%)	
DA	354	1,437	1,791	13.4%	15.1%
데이터 개발자	2,564	2,389	4,953	37.2%	10.7%
데이터 엔지니어	775	963	1,738	13.0%	9.5%
데이터 분석가	394	977	1,371	10.3%	14.0%
DBA	145	1,424	1,569	11.8%	8.1%
데이터 사이언티스트	158	338	496	3.7%	21.6%
데이터 컨설턴트	164	410	574	4.3%	10.3%
데이터 기획·마케터	311	534	845	6.3%	9.2%
전체	4,865	8,472	13,337	100.0%	10.9%

※ 인력 부족률 : 필요(부족)인력/(현 인력+필요(부족)인력)

(단위: 명, %)

나. 데이터직무 중 빅데이터 인력

1) 빅데이터 인력 현황

- 데이터직무 인력 중 빅데이터 관련 인력은 총 9,955명이며, 전체 데이터직무 인력의 9.1%를 차지한다.

[표12] 2017년 전 산업 내 데이터직무 인력 중 빅데이터 관련 인력 현황

구분	전 산업 내 데이터직무				전산업 내 데이터직무 중 빅데이터 인력 비중 (%)
	전체		빅데이터 기술 인력		
	규모(명)	비중(%)	규모(명)	비중(%)	
DA	10,071	9.2%	-	-	-
데이터 개발자	41,254	37.7%	2,908	29.2%	7.0%
데이터 엔지니어	16,634	15.2%	1,679	16.9%	10.1%
데이터 분석가	8,398	7.7%	1,108	11.1%	13.2%
DBA	17,863	16.3%	-	-	-
데이터 사이언티스트	1,803	1.7%	1,803	18.1%	100.0%
데이터 컨설턴트	5,004	4.6%	1,724	17.3%	34.5%
데이터 기획·마케터	8,293	7.6%	733	7.4%	8.8%
전체	109,320	100.0%	9,955	100.0%	9.1%

(단위: 명, %)

- ※ 빅데이터 관련 인력 : 데이터직무 인력 중 빅데이터 기술을 보유한 인력으로 데이터직무 인력 수에 포함됨
 ※ DA와 DBA는 조사 제외 대상이며, 데이터사이언티스트 직무는 100% 빅데이터 관련 인력으로 간주

2) 빅데이터 인력 수요

- 향후 3년 내 전 산업의 데이터직무 인력 수요 중 빅데이터 관련 인력 수요는 총 6,008명이며, 이 중 빅데이터 분석가⁴⁾(54.6%)와 빅데이터 개발자⁵⁾(44.7%)의 부족률이 가장 높게 나타났다.

[표13] 2017년 전 산업 내 데이터직무 인력 중 빅데이터 관련 인력 수요 현황

구분	전 산업 내 데이터직무 필요인력				인력 부족률 (%)
	전체		빅데이터 기술 필요인력		
	규모(명)	비중(%)	규모(명)	비중(%)	
DA	1,791	13.4%	-	-	-
데이터 개발자	4,953	37.2%	2,350	39.1%	44.7%
데이터 엔지니어	1,738	13.0%	907	15.1%	35.1%
데이터 분석가	1,371	10.3%	1,331	22.1%	54.6%
DBA	1,569	11.8%	-	-	-
데이터 사이언티스트	496	3.7%	496	8.3%	21.6%
데이터 컨설턴트	574	4.3%	509	8.5%	22.8%
데이터 기획·마케터	845	6.3%	415	6.9%	36.1%
전체	13,337	100.0%	6,008	100.0%	37.6%

(단위: 명, %)

- ※ 빅데이터 관련 인력 : 데이터직무 인력 중 빅데이터 기술을 보유한 인력으로 데이터직무 인력 수에 포함됨
 ※ DA와 DBA는 조사 제외 대상이며, 데이터사이언티스트 직무는 100% 빅데이터 관련 인력으로 간주
 ※ 인력 부족률 : 필요(부족)인력/(현 인력+필요(부족)인력)

4) 데이터분석가 중 빅데이터 기술 보유 인력
 5) 데이터개발자 중 빅데이터 기술 보유 인력

CONTENTS

제1부. 조사 개요

1.1 조사 개요	3
1) 조사목적	3
2) 법적근거	3
3) 조사연혁	3
4) 조사대상 및 범위	3
5) 조사내용	3
6) 조사주기와 기간	4
7) 조사방법	5
1.2 정의 및 분류체계	6
1) 데이터산업 정의	6
2) 데이터산업 현황조사 분류체계	7
1.3 조사대상	14
1) 모집단	14
2) 표본설계	16
1.4 모수 추정	20
1) 회수현황 및 이상치 처리	20
2) 단위 무응답 추정	22
3) 최종 추정치	23
4) 데이터산업 현황조사 이용 시 유의사항	25

제2부. 데이터산업 시장 규모

2.1 데이터산업 시장 규모	29
1) 전체 시장	29
2) 부문별 시장	31
2.2 데이터산업 직접매출 시장 규모	39
1) 전체 시장	39
2) 부문별 시장	41

제3부. 데이터산업 인력현황 및 수요

3.1 인력 현황	49
1) 데이터산업 인력 현황	49
2) 전 산업의 데이터직무 인력 현황	54
3.2 인력 수요	62
1) 데이터산업 인력 수요	62
2) 전 산업 데이터직무 인력 수요	66
3.3 데이터 자격증	70

제4부. IT 투자 및 정책 수요

4.1 IT 투자 현황	75
4.2 시장 환경 현황	77
4.3 정책 수요	78

제5부. 데이터 유통 현황

5.1 데이터 유통 현황	85
1) 데이터 거래 현황	85
2) 데이터 유통·활용 활성화를 위한 필요사항	92
5.2 데이터 보유 현황	94
1) 보유·제공 DB 현황	94
2) 보유·제공 DB 특징	96

그림목차

[그림1-1] 네이만 배분법에 의한 표본크기 결정 공식	16
[그림2-1] 데이터산업 시장 규모	29
[그림2-2] 데이터산업 부문별 시장 규모 비중	29
[그림2-3] 데이터산업 시장 전망	30
[그림2-4] 데이터솔루션 시장 규모	31
[그림2-5] 2017년 데이터솔루션 중분류별 시장 규모 비중	31
[그림2-6] 데이터구축/컨설팅 시장 규모	34
[그림2-7] 데이터구축 시장 규모	35
[그림2-8] 데이터컨설팅 시장 규모	35
[그림2-9] 데이터서비스 시장 규모	36
[그림2-10] 데이터서비스 중분류별 시장 규모 비중	36
[그림2-11] 데이터서비스 기반별 시장 규모	37
[그림2-12] 데이터산업 직접매출 시장 규모	39
[그림2-13] 데이터산업 부문별 직접매출 시장 규모 비중	39
[그림2-14] 데이터솔루션 직접매출 시장 규모	41
[그림2-15] 데이터구축/컨설팅 직접매출 시장 규모	43
[그림2-16] 데이터구축 직접매출 시장 규모	44
[그림2-17] 데이터컨설팅 직접매출 시장 규모	44
[그림2-18] 데이터서비스 직접매출 시장 규모	45
[그림2-19] 데이터서비스 중분류별 직접매출 시장 규모 비중	45
[그림2-20] 데이터서비스 기반별 직접매출 시장 규모	46
[그림3-1] 데이터산업 인력 현황	49
[그림3-2] 데이터산업 내 데이터직무별 인력 비중	51
[그림3-3] 전 산업 내 데이터직무 비중	54
[그림3-4] 2017년 전년대비 전 산업 내 데이터직무별 증감률	56
[그림3-5] 2017년 데이터산업 내 데이터직무 인력 부족률	63
[그림3-6] 2017년 전 산업 내 데이터직무별 인력 부족률	67
[그림4-1] 데이터산업 전체 세부 투자비중	75
[그림4-2] 데이터산업 정책 수요	78
[그림4-3] 데이터산업 필요 정보	79
[그림5-1] 데이터 거래 경험 여부	85
[그림5-2] 데이터 거래 목적	86
[그림5-3] 구매처별 거래 비중	87
[그림5-4] 판매처별 거래 비중	89
[그림5-5] 데이터 거래 시 애로사항	91

표목차

[표1-1] 조사 연혁	3
[표1-2] 데이터산업 현황조사 세부 조사항목	4
[표1-3] 데이터산업 범위	6
[표1-4] 2017년 데이터산업분류 및 한국표준산업분류 비교	7
[표1-5] 데이터산업 분류체계 2017년-2016년 비교	8
[표1-6] 데이터솔루션 시장 정의	10
[표1-7] 데이터구축 및 컨설팅 시장 정의	12
[표1-8] 데이터산업 현황조사 조사 모집단 구축 내용	14
[표1-9] 모집단 분포	15
[표1-10] 할당방법별 표본의 크기 및 상대표준오차	17
[표1-11] 표본 분포	19
[표1-12] 조사기업 분포	20
[표1-13] 데이터산업 현황조사 회수율	20
[표1-14] 항목 간 논리 점검 예시	21
[표1-15] 추정 대상 무응답 사업체 수	22
[표1-16] 데이터산업 매출액 상대표준오차, 신뢰수준	24
[표1-17] 데이터산업 직접매출액 상대표준오차, 신뢰수준	24
[표1-18] 데이터직무 인력 상대표준오차, 신뢰수준	24
[표2-1] 데이터산업 시장 규모 총괄	30
[표2-2] 데이터산업 시장 전망	30
[표2-3] 데이터솔루션 중분류별 시장 규모	32
[표2-4] 데이터솔루션 영역별 시장 규모	32
[표2-5] 데이터솔루션 업종별 매출 비중	33
[표2-6] 데이터 구축/컨설팅 시장 규모	34
[표2-7] 데이터서비스 중분류별 시장 규모	37
[표2-8] 데이터서비스 기반별 시장 규모	38
[표2-9] 데이터산업 부문별 직접매출 시장 규모	40
[표2-10] 데이터솔루션 중분류별 직접매출 시장 규모	41
[표2-11] 데이터솔루션 영역별 직접매출 시장 규모	42
[표2-12] 데이터 구축/컨설팅 직접매출 시장 규모	43
[표2-13] 데이터서비스 중분류별 직접매출 시장 규모	46
[표2-14] 데이터서비스 기반별 직접매출 시장 규모	46
[표3-1] 데이터산업 인력 현황	49
[표3-2] 데이터산업 내 부문별 데이터직무 인력 현황	50
[표3-3] 데이터직무 분류표	50
[표3-4] 2017년 데이터산업 내 데이터직무별 인력 현황	51
[표3-5] 데이터산업 내 데이터직무 중 빅데이터 관련 인력 현황	52
[표3-6] 데이터산업 내 데이터직무 중 빅데이터 관련 인력 비중	53

표목차

[표3-7] 데이터산업 내 부문별 빅데이터 관련 인력 현황	53
[표3-8] 전 산업 내 데이터직무 인력 현황	54
[표3-9] 2017년 전 산업 내 데이터직무별 인력 현황	55
[표3-10] 2015년-2017년 전 산업 내 데이터직무별 인력 현황 비교	56
[표3-11] 데이터기술 등급 기준	57
[표3-12] 전 산업 내 데이터직무 기술등급별 인력 현황	58
[표3-13] 전 산업 내 데이터직무 성별 인력 현황	59
[표3-14] 전 산업 내 데이터직무 중 빅데이터 관련 인력 현황	60
[표3-15] 전 산업 내 데이터직무 중 빅데이터 관련 인력 비중	61
[표3-16] 전 산업 내 부문별 데이터직무 중 빅데이터 관련 인력 현황	61
[표3-17] 2017년 데이터산업 내 데이터직무별 필요인력	62
[표3-18] 2017년 데이터산업 내 데이터직무 인력 부족률	63
[표3-19] 데이터산업 내 데이터직무 중 빅데이터 관련 필요인력	64
[표3-20] 데이터산업 내 부문별 빅데이터 관련 필요인력	65
[표3-21] 2017년 데이터산업 내 데이터직무 중 빅데이터 관련 인력 부족률	65
[표3-22] 향후 3년 내 전 산업 내 데이터직무별 필요인력	66
[표3-23] 2017년 전 산업 내 데이터직무별 인력 부족률	67
[표3-24] 전 산업 내 데이터직무 중 빅데이터 관련 필요인력	68
[표3-25] 전 산업 내 부문별 데이터직무 중 빅데이터 관련 필요인력	69
[표3-26] 2017년 전 산업 내 데이터직무 중 빅데이터 인력 부족률	69
[표3-27] 데이터산업의 데이터 전문 인력 우대 자격증	70
[표3-28] 데이터산업의 연도별 우대 자격증 비교	70
[표3-29] 전 산업의 데이터 전문 인력 우대 자격증	71
[표3-30] 전 산업의 데이터 전문 인력 우대사항	71
[표3-31] 전 산업의 자격증별 데이터 전문 인력 우대사항	72
[표4-1] 데이터 관련 사업 투자비중	75
[표4-2] 데이터산업 분야별 세부 투자비중	76
[표4-3] 데이터 관련 사업 주요 고객	77
[표4-4] 데이터 관련 사업 주요 경쟁국	77
[표4-5] 데이터산업 부문별 정책 수요	79
[표4-6] 데이터산업 부문별 필요 정보	80
[표4-7] 데이터사업 수행 시 애로사항	80
[표4-8] 기술관련 자산 및 데이터베이스의 매출 기여 비중	81
[표5-1] 데이터서비스 분야별 데이터 거래 경험 여부	85
[표5-2] 데이터 거래 경험 없는 이유	86
[표5-3] 데이터서비스 분야별 구매처별 거래 비중	87
[표5-4] 구매 거래 가격 형태	88

표목차

[표5-5] 데이터서비스 분야별 구매처 유형별 거래 가격 형태	88
[표5-6] 데이터서비스 분야별 판매처별 거래 비중	89
[표5-7] 판매 거래 가격 형태	90
[표5-8] 데이터서비스 분야별 판매처 유형별 거래 가격 형태	90
[표5-9] 데이터서비스 분야별 데이터 거래 시 애로사항	91
[표5-10] 데이터 유통·활용 활성화 필요지원 사항	92
[표5-11] 거래경험 유무별 데이터 유통·활용 활성화 필요지원 사항	92
[표5-12] 개인정보가 포함된 데이터 유통·활용 필요지원 사항	93
[표5-13] 거래경험 유무별 개인정보가 포함된 데이터 유통·활용 필요지원 사항	93
[표5-14] DB 최초 개발 시기	94
[표5-15] DB 개발 완료 소요기간 및 구축 비용	94
[표5-16] DB 개발 투입인원	94
[표5-17] 갱신 주기	95
[표5-18] 갱신 소요비용	95
[표5-19] 갱신 투입인력	95
[표5-20] DB 수록데이터 유형	96
[표5-21] 데이터 상품 성격	96
[표5-22] DB 유통 방식	96

The background features a stylized city skyline in shades of blue and grey on the left side. The rest of the page is composed of large, overlapping geometric shapes in various shades of blue, creating a modern, abstract design.

제1부. 조사 방법

1.1 조사 개요

1) 조사목적

- 본 조사의 목적은 급변하는 데이터산업 환경 변화를 파악하여 체계적인 데이터산업 육성 지원 정책 수립을 위한 기초 통계를 제공하는 것이다.

2) 법적근거

- 정부승인통계 (승인번호 127004)

3) 조사연혁

[표1-1] 조사 연혁

연도	내용
1995년 ~ 2002년	데이터베이스 산업 현황 및 이용 실태 조사
2003년 ~ 2014년	데이터베이스 산업 현황 조사
2015년 ~	데이터산업 현황 조사

4) 조사대상 및 범위

- 본 조사는 전국에 소재한 데이터산업 영위 사업체를 대상으로 한다.
- 데이터산업 영위 사업체는 데이터의 생산, 수집, 처리, 분석, 유통, 활용 등을 통해 가치를 창출하는 상품과 서비스를 생산·제공하는 1인 이상 사업체를 말한다.

5) 조사내용

- 데이터산업은 비즈니스 성격에 따라 데이터솔루션, 데이터서비스, 데이터구축, 데이터컨설팅 4개 영역으로 구분되고 있으며, 대상에 따라 조사내용을 차별화하였다.
- 2017년 조사에서는 2016년 조사표 문항을 기본 항목으로 구성하여 조사의 연속성을 유지하였다.

[표1-2] 데이터산업 현황조사 세부 조사항목

구분	조사 내용	대상
매출 현황	- 전체 데이터 매출 - 세부 영역별, 업종별 매출	데이터 기업 전체
인력 현황	- 전체 종사자 수 - 데이터직무별 현재인력 수 - 데이터직무별 필요인력 수 - 전문 인력 우대자격증 및 우대사항	
투자 현황 및 시장 환경	- 데이터산업 투자비중 - 세부 투자금액 비중 - 주요 고객 - 주요 경쟁국 및 경쟁사	
데이터 유통 현황	- 데이터 거래 경험, 거래 경험이 없는 이유 - 구매처·판매처별 거래 비중 - 데이터 거래 목적, 데이터 거래 애로사항 - 데이터 유통 활성화 지원 사항 - 개인데이터 유통 시 필요 지원 사항	데이터 서비스 기업
데이터 보유 현황	- DB명, DB설명, 서비스개시일 - 구축 소요기간, 구축비용, 구축 투입인력 수 - 데이터 총량 또는 레코드 수 - 갱신 주기, 갱신 소요비용, 갱신 투입인력 수 - DB 데이터 유형, DB 유통 방식	
업계 의견 수렴	- 애로사항 - 필요한 정책 지원 사항 - 데이터사업 추진 시 필요 정보	데이터 기업 전체
	기술관련 자산 및 데이터베이스의 매출 기여도	데이터 서비스 기업

6) 조사주기와 기간

- 조사주기 : 1년
- 조사기준일 : 2016. 12. 31. 및 2017. 6. 30.⁶⁾
- 조사실사 기간 : 2017. 10. ~ 2017. 12.

6) 매출액 조사 시점을 2016년 확정치와 2017년 상반기 매출을 모두 조사함

7) 조사방법

○ 본 조사에 적합한 실사팀을 구성

- 조사목적, 조사대상, 설문지에 대한 전반적인 실사진행에 대해 논의하고 실사지침을 공유하였다.

○ 모집단 리스트 확보 및 정리

- 「2016년 데이터산업 현황조사」를 기초로 하여 관련 리스트를 수집 후 모집단 리스트를 확정하였다. 각 분야별로 기업규모를 고려하여 표본을 설계하고 표본과 대체표본을 구분하였다.

○ 전문조사원 선발 및 교육

- 기업체(사업체)조사 경력 3년 이상의 전문 조사원 중 평가점수가 높은 조사원을 우선 선발하고, 실사지침을 바탕으로 조사에 대한 전반적인 교육을 실시한 후 조사에 투입하였다.

○ 조사 진행

- 1차적으로 조사원이 조사대상 사업체에 전화조사를 수행하며 조사의 목적과 취지를 설명하였다.
- 조사 협조 공문 활용, 조사의 기대효과와 중요한 안내 등 다양한 조사 동기 유발 방법을 실시하여 조사 참여율과 자료의 회수율을 높이고자 노력하였다.

○ 자료검수 및 통계처리

- 조사의 신뢰성 확보를 위해 사전에 작성된 에디팅 가이드에 따라 자료검수를 실시하였다. 부실 자료에 대한 재확인과 보완을 거친 후 최종 유효자료에 대해 통계처리를 하였다.

1.2 정의 및 분류체계

1) 데이터산업 정의

- 데이터산업은 '데이터의 생산·수집·처리·분석·유통·활용 등을 통해 가치를 창출하는 상품과 서비스를 생산·제공하는 산업'으로 정의한다. 데이터의 생명주기 또는 가치사슬 상에 나타난 데이터의 생산, 수집, 분석, 유통 및 활용 등 제반 활동을 포함하여 데이터로부터 가치가 창출되는 일련의 모든 과정, 그리고 이와 연관된 활동들을 포함하고 있다.
- 이러한 정의에 따라 데이터산업의 비즈니스 유형은 크게 데이터 관련 제품을 판매하거나 기술을 제공하는 데이터솔루션, 데이터구축, 데이터컨설팅 비즈니스와 데이터를 기반으로 서비스를 제공하는 데이터서비스 비즈니스로 구분할 수 있다. 데이터산업의 시장 규모는 이러한 데이터 비즈니스를 영위하는 사업체들의 데이터 관련 매출을 조사하여 추정한다. 데이터 관련 매출에는 직접매출 외에도 데이터를 매개로 발생하는 간접매출, 가령 광고매출 등이 있을 수 있으며, 주로 데이터서비스 분야에서 발생한다.
- 본 보고서에서는 전통적인 DB산업을 아우르는 기존 조사 범위를 적용해 데이터 관련 간접매출을 포함하는 광의의 데이터산업 시장 규모(이하 '데이터산업 시장 규모')와 데이터 직접매출 중심의 협의의 데이터산업 시장 규모(이하 '데이터산업 직접매출 시장 규모')를 함께 추정하였다.

[표1-3] 데이터산업 범위

구분	범위
데이터 솔루션	• DBMS, DBMS관리, 데이터 모델링, 분석·시각화, 검색엔진, 품질 등 관련 솔루션 제품으로 비즈니스를 영위하는 기업 • 라이선스, 유지보수, 커스터마이징(개발)에서 매출 발생
데이터 서비스	• 데이터를 활용해 정보제공, 데이터 거래, 분석 결과 정보 등을 온·오프라인(모바일, 앱 등 포함)으로 제공하면서 데이터 이용료/수수료 또는 광고료 등으로 비즈니스를 영위하거나, 마케팅을 목적으로 데이터를 수집·가공하여 판매하는 기업(데이터 브로커)
데이터 구축	• DB설계, 데이터 이행 등을 포함한 DB시스템 구축, 문서·음성·영상 등의 데이터를 DB로 변환·정비하는 데이터 처리, 데이터 외부 제공을 위한 API, LOD 구축, DW, Data Lake 등의 데이터 구축 등의 비즈니스를 영위하는 기업(데이터/DB 관련 SI·IT아웃소싱 포함)
데이터 컨설팅	• 데이터 거버넌스, 품질, 데이터 설계 및 데이터 활용 등 데이터 관련 기획 및 컨설팅 비즈니스를 영위하는 기업

○ 2017년 데이터산업분류와 한국표준산업분류 비교는 다음과 같다.

[표1-4] 2017년 데이터산업분류 및 한국표준산업분류 비교

2017년 데이터산업분류			한국표준산업분류 9차 개정
대분류	중분류	소분류	
데이터 솔루션	데이터 수집	데이터 검색 로그데이터 수집 웹데이터 수집	J58222 응용소프트웨어 개발 및 공급업 J62010 컴퓨터 프로그래밍 서비스업
	데이터 설계	데이터 아키텍처 데이터 모델링	
	DBMS	디스크 DBMS 인메모리 DBMS DBMS 어플라이언스 기타 DBMS	
	데이터 관리	DB 운영 관리 DB 성능 관리 DB 보안 관리 데이터 분산 처리 데이터 통합/흐름 관리	
	데이터 품질	데이터 품질 관리 메타데이터 관리 마스터데이터 관리	
	데이터 분석	정형 데이터 분석 비정형 데이터 분석 실시간 데이터 분석 데이터 시각화 분석	
	데이터 플랫폼	데이터 플랫폼	
데이터 구축/컨설팅	구축	DB설계·구축 데이터 이행·처리 기계처리형 데이터 구축	J58221 시스템 소프트웨어 개발 및 공급업 J62021 컴퓨터시스템 통합 자문 및 구축 서비스업 J62090 기타 정보기술 및 컴퓨터운영 관련 서비스업 J63111 자료처리업
	컨설팅	데이터 설계 데이터 품질 데이터 관리 성능개선 데이터 거버넌스 체계 데이터 활용	
데이터 서비스	데이터 거래	데이터 마켓 데이터 신디케이션	J63120 포털 및 기타 인터넷 정보매개 서비스업 J63910 뉴스제공업 J63991 데이터베이스 및 온라인정보 제공업 J63999 그 외 기타 정보서비스업
	정보 제공	교육/훈련 신용/재무 통계 경영/비즈니스 취업/창업 뉴스 포털 행정/법률 학술 문화/예술 생활	
	데이터 분석 제공	환경분석서비스 예측분석서비스 심사분석서비스 기타 데이터분석서비스	

2) 데이터산업 현황조사 분류체계

- 2017년 데이터산업 분류체계는 아래 표와 같이 총 3개 대분류, 12개 중분류, 47개 소분류로 구성된다.(조사는 중분류 단위로 실시)
- 2017년 조사과정에서 2016년 데이터산업 분류에 대한 업계 의견수렴을 한 결과, 컨설팅의 소분류인 '데이터 활용'을 '데이터 분석'으로 용어를 변경하였다.

[표1-5] 데이터산업 분류체계 2017년-2016년 비교

2017년 데이터산업			비고	2016년 데이터 산업		
대분류	중분류	소분류		대분류	중분류	소분류
데이터 솔루션	데이터 수집	데이터 검색		데이터 솔루션	데이터 수집	데이터 검색
		로그데이터 수집				로그데이터 수집
		웹데이터 수집				웹데이터 수집
	데이터 설계	데이터 아키텍처			데이터 설계	데이터 아키텍처
		데이터 모델링				데이터 모델링
	DBMS	디스크 DBMS			DBMS	디스크 DBMS
		인메모리 DBMS				인메모리 DBMS
		DBMS 어플라이언스				DBMS 어플라이언스
		기타 DBMS				기타 DBMS
	데이터 관리	DB 운영 관리			데이터 관리	데이터 관리
		DB 성능 관리				DB 성능 관리
		DB 보안 관리				DB 보안 관리
		데이터 분산 처리				데이터 분산 처리
		데이터 통합·흐름 관리				데이터 통합·흐름 관리
	데이터 품질	데이터 품질 관리			데이터 품질	데이터 품질 관리
		메타데이터 관리				메타데이터 관리
		마스터데이터 관리				마스터데이터 관리
	데이터 분석	정형 데이터 분석			데이터 분석	정형 데이터 분석
		비정형 데이터 분석				비정형 데이터 분석
		실시간 데이터 분석				실시간 데이터 분석
		데이터 시각화 분석				데이터 시각화 분석
	데이터 플랫폼	데이터 플랫폼			데이터 플랫폼	데이터 플랫폼

(계속)

2017년 데이터산업			비고	2016년 데이터 산업		
대분류	중분류	소분류		대분류	중분류	소분류
데이터 구축/ 컨설팅	구축	DB설계·구축 데이터 이행·처리 기계처리형 데이터구축	용어변경	데이터 구축/ 컨설팅	구축	DB설계·구축 데이터 이행·처리 기계처리형 데이터구축
	컨설팅	데이터 설계 데이터 품질 데이터 관리 성능개선 데이터 거버넌스 체계 데이터 분석			컨설팅	데이터 설계 데이터 품질 데이터 관리 성능개선 데이터 거버넌스 체계 데이터 활용
데이터 서비스	데이터 거래	데이터 마켓 데이터 신디케이션		데이터 서비스	데이터 거래	데이터 마켓 데이터 신디케이션
	정보 제공	교육/훈련 신용/재무 통계 경영/비즈니스 취업/창업 뉴스 포털 행정/법률 학술 문화/예술 생활			정보 제공	교육/훈련 신용/재무 통계 경영/비즈니스 취업/창업 뉴스 포털 행정/법률 학술 문화/예술 생활
	데이터 분석 제공	환경분석서비스 예측분석서비스 심사분석서비스 기타 데이터분석서비스			데이터 분석 제공	환경분석서비스 예측분석서비스 심사분석서비스 기타 데이터분석서비스

○ 데이터산업 분류별 상세한 정의 및 관련 예시는 다음과 같다.

[표1-6] 데이터솔루션 시장 정의

대분류	중분류	소분류	정의	관련 예시
데이터 솔루션	데이터 수집	데이터 검색	- 사용자가 원하는 데이터를 찾을 수 있도록 지원하는 솔루션 - 사용자가 필요로 하는 데이터를 다양한 검색조건을 지정하여 데이터세트를 만들고 활용할 수 있도록 지원하는 솔루션	검색엔진, Query
		로그 데이터 수집	다양한 시스템으로부터 발생하는 로그를 실시간으로 수집하고, 이중 간 발생하는 다양한 정보의 형태를 정형화하여 통합 관리하는 솔루션	공정로그 수집 솔루션, 고객의견 (VOC) 로그 수집 솔루션
		웹데이터 수집	조직적, 자동화된 방법으로 인터넷 상에 존재하는 웹 문서들을 추적하여 필요한 데이터를 수집하는 솔루션	웹 크롤러
	데이터 설계	데이터 아키텍처	기업 또는 기관의 전사 차원에서 데이터 현황을 파악하고 관리할 수 있는 솔루션	데이터 참조모델 생성 및 활용, 다중 DBMS 저장소, 데이터 아키텍처 관리 솔루션 등
		데이터 모델링	데이터베이스 구축을 위해 데이터 모델링을 수행하고 모델링 결과에 따라 데이터베이스 물리설계도를 생성하는 솔루션	데이터 모델링 솔루션 등
	DBMS	디스크 DBMS	기존에 주로 이용하고 있는 DBMS로서, 데이터를 별도의 물리적 저장장치에 보관하는 DBMS	전통적 DBMS
		인메모리 DBMS	새롭게 등장한 DBMS로서, 활용하고자 하는 데이터를 인메모리(in-memory)에 올려놓고 운영함으로써 성능이 개선된 DBMS	실시간 DBMS, 인메모리 기반 DBMS 등
		DBMS 어플라이언스	DBMS가 운영될 하드웨어와 저장장치를 DBMS와 통합하여 하나의 하드웨어로 제공하는 솔루션 또는 서버	DBMS가 내재화된 서버
		기타 DBMS	전통적인 DBMS, 인메모리 DBMS, DBMS 어플라이언스에 포함되지 않는 DBMS	클라우드 DBMS, 하이브리드 DBMS, 기타DBMS 등
	데이터 관리	DB 운영 관리	- DB 운영 단계에서 적용하는 도구로서, 주로 DB 형상관리(configuration management)에 적용되는 솔루션 - 비즈니스 연속성을 위한 데이터의 관리를 위한 백업/복구 소프트웨어	DB 운영 솔루션, 백업 및 복구, 데이터 중복 제거 등
		DB 성능 관리	DB성능을 모니터링하고 장애 및 지연 요인을 발견하고 해결하기 위한 솔루션	대용량 데이터 처리, 자동진단(장애 분석), 자동튜닝, 능력(CAPA) 산정, DB성능 모니터링 솔루션
		DB 보안 관리	DB의 유출, 개량, 파괴, 접근 등으로부터 DB의 보안을 위한 방화벽, 침입탐지/방지, 서버 보안 등 데이터베이스 접근 통제, 암호화, 보안감사를 수행하는 소프트웨어	멀티DB 암호화 및 접근 통제 연계, 클라우드DB 보안, 능동적DB보안, 대규모 스트림 데이터 암호화, 데이터 익명화, DB포렌식(Forensic) 감사 등 DB 암호화 및 접근통제 솔루션
		데이터 분산 처리	기존의 데이터 처리 기술로는 효율적으로 처리하기 어려운 대용량의 비구조화 데이터 처리를 위하여 대량의 하드웨어를 이용하여 분산 저장(분산파일 시스템, 분산 스토리지 시스템), 처리(맵리듀스)하는 솔루션	클라우드 컴퓨팅, 하둡
		데이터 통합·흐름 관리	운영 데이터베이스로부터 데이터를 추출하여 데이터웨어하우스로 정제 및 가공하여 적재하는 솔루션, 기업 및 기관 내부의 데이터 흐름을 자동화하는 솔루션	EAI(Enterprise Application Interface), ETL(Extract, Transform, Load)

(계속)

대분류	중분류	소분류	정의	관련 예시
데이터 솔루션	데이터 품질	데이터 품질 관리	운영 데이터베이스에 보관되고 있는 데이터의 품질을 측정하고 평가하며, 변경 영향을 분석하는 솔루션	실시간 데이터 품질 검측, 비정형데이터 쿼리 매칭 등 데이터 품질 관리 솔루션
		메타 데이터 관리	운영 데이터베이스의 데이터 정의와 표준 등 메타데이터를 관리하고 데이터 변경을 통제하는 솔루션	데이터 표준 자동화, 대용량 분산 메타데이터 관리, 영향도 분석 등 메타데이터 관리 솔루션
		마스터 데이터 관리	기업 활동의 기준이 되는 핵심 데이터를 식별하고 전사적으로 일관되게 사용하기 위해 적용하는 솔루션	플랫폼형 마스터데이터 관리 기술, 개발 프레임워크 일체형 마스터데이터 관리 도구, MDM과 빅데이터 결합, 마스터데이터 추출 및 분리 기술 등 Master Data Management Solution
데이터 분석	정형 데이터 분석	정형 데이터 분석	- DBMS의 데이터를 사용자의 이용목적에 맞도록 간편하게 보고서를 생성 및 활용 할 수 있도록 지원하는 솔루션(데이터리포팅) - 미리 정의된 관점에 따라 사용자가 기준을 유연하게 변경하면서 데이터를 분석할 수 있도록 지원하는 솔루션(OLAP) - 데이터 집합을 통계적으로 분석하여 의미 있는 인사이트를 찾아내는 과정을 지원하는 솔루션 등 (데이터마이닝)	Reporting 솔루션, OLAP(Online Analytical Processing), 통계분석패키지
		비정형 데이터 분석	- 웹사이트 데이터를 추출 및 분석하여 기술트렌드 및 기업 평판 등을 분석하기 위해 적용하는 솔루션 (웹크롤링) - 주로 문서 및 웹사이트의 데이터에서 키워드를 찾아내고 빈도 및 연관관계 등을 분석하는 솔루션 등 (텍스트마이닝) - 신문기사, 논문, 특허 등의 동향정보로부터 주요 개체 및 개체와 관련된 사실정보를 정형화된 형태로 추출하여 분석에 활용하는 솔루션(정보추출) - 음성 및 영상 등의 데이터로부터 의미 있는 데이터를 추출하여 분석하는 솔루션 - 소셜미디어 내의 인맥정보를 이용하여 사회관계망을 분석하고, 포스팅 되는 내용을 통해 최신 이슈와 트렌드를 분석하는 솔루션	웹 크롤링(Web Crawling), 텍스트 마이닝, 비디오데이터 가공 솔루션, 음성 데이터 가공 솔루션, 트위터, 페이스북 등의 API 연계 솔루션 (Social Analytics 솔루션)
	실시간 데이터 분석	실시간 데이터 분석	데이터 분석을 통해 위험 신호 또는 징후를 감지하여 위기나 재난에 대처할 수 있도록 예측하는 솔루션	CEP(Complex Event Processing), BAM(Business Activity Monitoring)
		데이터 시각화 분석	데이터 세트의 의미를 시각적으로 표현하거나 탐색하는 과정을 지원하는 솔루션	시각화 및 비주얼 분석 솔루션 (Visual Analytics 솔루션)
	데이터 플랫폼	데이터 플랫폼	개인, 사업자, 기관 등이 보유한 원천데이터를 변환, 가공, 정제, 발행(publishing), 저장, 질의 등이 가능하도록 하는 솔루션 혹은 기능들의 집합	빅데이터 플랫폼 등

[표1-7] 데이터구축 및 컨설팅 시장 정의

대분류	중분류	소분류	정의	관련 예시
데이터 구축/ 컨설팅	구축	DB설계· 구축	- 업무요건을 충족시킬 수 있는 데이터베이스의 구조와 형태 및 속성을 정의하는 서비스 - 정형 또는 비정형의 대량 데이터를 가공전 raw data로 저장하는 리포지토리나 사용자가 쉽게 접근할 수 있는 분석용(의사결정) 데이터베이스 등을 구축하는 서비스	물리 DB 설계, 데이터웨어하우스(DW), 데이터마트(Data Mart), 데이터레이크(Data Lake) 등 구축
		데이터 이행· 처리	- 시스템 재구축으로 인해 기존에 운영 중인 데이터베이스에 보관 중인 데이터를 점검하고 정제하여 새롭게 개발하고 있는 데이터베이스로 이행하는 서비스 - 일반문서, 음성, 영상 등의 자료를 데이터베이스에 보관 할 수 있도록 형태를 정비하거나 변환하는 서비스 - 다양한 데이터를 주어진 형식에 맞추어 구축하는 서비스	차세대 시스템 구축, 데이터 변환 등
		기계 처리형 데이터 구축	- 기업 또는 기관이 보유하고 있는 데이터를 외부에 제공 및 활용할 수 있도록 API, LOD(Linked Open Data) 등을 개발, 구축하는 서비스 - 기존에 존재하지 않았던 데이터를 특별 목적을 위해 데이터화(datafication) 하는 것 포함	API 구축, LOD 구축 등 machine processable 데이터 구축
컨설팅	컨설팅	데이터 설계	- 전사 차원에서 데이터 요건을 분석하여, 단위 데이터 저장소를 정의하고 데이터 저장소들의 연관 관계를 정의하는 서비스 - 데이터 아키텍처를 기반으로 특정 영역의 업무 요건을 보다 상세하고 완전하게 분석하여 데이터베이스를 구축할 수 있는 수준으로 설계하는 서비스 - 데이터 설계를 기획하거나 가이드, 리딩 또는 지원하는 서비스	Data Architecture Planning, Analytics Strategy Planning, Data Modeling, Data Reference Model 등
		데이터 품질	운영하고 있는 데이터베이스에 포함된 실제 데이터의 품질을 점검하고, 개선점을 제시하는 서비스. 데이터 정의, 데이터 값, 데이터 연관관계 등의 품질 영역이 있음	데이터 품질 심사 및 인증, 데이터 품질 진단 등
		데이터 관리 성능개선	운영하고 있는 데이터베이스의 활용 과정에서 온라인 및 배치 작업의 병목 지점을 발견하고 이를 개선함으로써 요구하는 서비스 수준을 달성시키기 위한 서비스	DB튜닝, DB성능 진단 등
		데이터 거버넌스 체계	데이터 아키텍처로부터 데이터 모델, 물리 데이터베이스 설계안 등이 현재 운용 중인 데이터베이스의 모습과 동시성을 유지하면서, 변경 요인을 효과적으로 반영할 수 있는 조직, 역할, 프로세스 등을 설계하는 서비스	데이터 거버넌스 진단, 데이터 거버넌스 체계 수립 등
		데이터 분석	내부 및 외부 데이터를 활용하여, 기업 또는 기관에서 필요로 하는 분석 기법을 개발하고, 테스트하여, 분석 모듈을 개발할 수 있도록 컨설팅해주는 서비스	추천엔진 개발, FDS(Fraud Detection System) 개발 등

(계속)

대분류	중분류	소분류	정의	관련 예시
데이터 서비스	데이터 거래	데이터 마켓	• (데이터 중개) 데이터 유통 플랫폼을 통해 데이터 보유자(보유기업)와 수요자(수요기업)간 원천데이터를 중개하는 서비스	API스토어, 데이터허브, 데이터스토어, 데이터 오픈마켓, 데이터중개 등
			• (데이터 판매) 온오프라인, API, 자사 유통시스템 등을 이용해 기업 간(B2B)에 원천데이터를 판매하는 서비스	데이터/DB 판매 등
		데이터 신디케이션	• 다양한 데이터를 바탕으로 고객이 원하는 형태로 또는 특정 주제에 따라 구성하여 제공하는 맞춤형 데이터 서비스	데이터 융합
	정보 제공	교육/훈련	• 자격증, 시험, 유학, 연수 등의 교육/훈련 관련 정보 서비스	
		신용/재무	• 기업일반, 기업·개인 신용, 재무·회계, 보험/증권/주식/환율 등의 정보서비스	
		통계	• 시장정보, 리서치정보 등 통계, 시장동향 정보서비스	
		경영/비즈니스	• 쇼핑·고객, 제품·상품·부품, 물류·운송, 경매·입찰, 부동산 정보서비스	
		취업/창업	• 취업(구인구직), 창업 관련 정보서비스	
		뉴스	• 뉴스제공, 뉴스사진제공 등 뉴스를 제공하는 서비스	
		포털	• 다양한 정보를 검색해 접근할 수 있도록 하는 포털 형식의 정보서비스	
		행정/법률	• 법률·법규·판례, 행정·민원, 조세·세법, 지식재산권 관련 정보서비스	
		학술	• 인문사회과학, 자연과학, 기술과학, 의학, 참고자료 등 관련분야 전문정보 및 논문 정보서비스	
		문화/예술	• 문학·만화, 미술, 사진, 음악, 공연예술, 매체예술, 전통문화 등 문화예술 관련 정보서비스	
		생활	• 관광·여행, 건강·보건, 교통, 지도·지리, 날씨·기상, 인물, 도서출판, 종교, 결혼·장례, 여성·육아, 오락·게임, 스포츠·레저 등 생활과 밀접한 정보를 제공하는 정보서비스	
	데이터 분석 제공	환경분석 서비스	• 소셜데이터, 오픈데이터 등 외부데이터와 자체보유하고 있는 데이터를 연계하여 주요 환경 변화 트렌드를 정보로 생성하여 제공하는 서비스	Horizontal Scanning, Reference Engine, 부정사용방지 또는 이상거래 탐지시스템(FDS; Fraud Detection System), SNS 대화 감정분석 서비스, 여론분석서비스, 상권분석서비스 등
		예측분석 서비스	• 서비스 및 상품추천, 위험 사전 예보 등의 분석을 대행하여 수행하고 그 결과를 제공하는 서비스	
		심사분석 서비스	• 개인신용등급 등 금융 및 상거래를 위해 필요한 심사 분석을 대행하여 수행하고 그 결과를 제공하는 서비스 등	
		기타 데이터 분석 서비스	• 그 외 대량의 다양한 데이터를 분석하여 유의미한 값을 도출, 제공하는 서비스	

1.3 조사대상

1) 모집단

가. 모집단 정의

- 데이터산업이란, 데이터의 생산, 수집, 처리, 분석, 유통, 활용 등을 통해 가치를 창출하는 상품과 서비스를 생산·제공하는 산업이다.

나. 모집단 정비

- 한국표준산업분류의 데이터산업 관련 세세분류를 도출 후 해당 사업체의 데이터산업 영위여부를 파악하는 '모집단 정비 사전조사'를 통해 모집단 사업체를 구성하였다.⁷⁾ 과년도 데이터산업 현황 조사에서 확인된 사업체와 정보통신정책연구원(KISDI)이 보유한 데이터 관련 사업체 리스트를 통합하여 조사대상 사업체를 수집하였다.
- 1차적으로 확보된 모집단 리스트에서 중복된 사업체를 제거하고, 사업자등록번호가 있는 경우에는 국세청 홈페이지에서 휴·폐업 등 경영상태, 업종, 주소 변경 등 사업체 기본정보를 확인한 후, 한국데이터진흥원의 검토를 거쳐 모집단을 정제하였다.

[표1-8] 데이터산업 현황조사 조사 모집단 구축 내용

구분	내용
통계청 모집단 리스트	• 2015년 기준 전국사업체조사(통계청) (58221) 시스템 소프트웨어 개발 및 공급업 (58222) 응용소프트웨어 개발 및 공급업 (62021) 컴퓨터시스템 통합 자문 및 구축 서비스업 (62090) 기타 정보기술 및 컴퓨터운영 관련 서비스업 (63111) 자료처리업 (63120) 포털 및 기타 인터넷정보매개서비스 (63910) 뉴스 제공업 (63991) 데이터베이스 및 온라인 정보 제공업 등 관련 기업 등 (63999) 그 외 기타 정보 서비스업
모집단 확인	• 중복 확인 • 휴·폐업, 청산·해산 등 경영상태 확인 • 업종, 주소 변경 등 기본정보 확인
최종 모집단 구축	• 데이터솔루션, 서비스, 구축, 컨설팅 최종 모집단 선정

7) 통계청 전국사업체조사(2015년 기준)를 통해 데이터산업 모집단을 선정하는데 참조하여 1차 모집단 구축

다. 조사 모집단

- 2017년 데이터산업 조사 모집단은 총 7,133개이다. 세부 산업분야별 모집단은 데이터솔루션이 1,608개, 데이터구축 1,594개, 데이터컨설팅 1,004개, 데이터서비스는 2,927개이다.
- 종사자 규모별로는 1~9인 사업체가 5,242개로 가장 많으며, 10~49인 사업체는 1,293개, 1,000인 이상 사업체는 29개로 나타났다.

[표1-9] 모집단 분포

업종		종사자 규모						총합
		1-9인	10-49인	50-99인	100-299인	300-999인	1,000인 이상	
데이터 솔루션	1. 데이터 수집	274	47	13	5	2		341
	2. 데이터 설계	171	36	10	4	2	1	224
	3. DBMS	109	14	4	3	4	1	135
	4. 데이터 관리	302	115	30	21	2	2	472
	5. 데이터 품질	136	15	5	2			158
	6. 데이터 분석	125	40	8	7	2	1	183
	7. 데이터 플랫폼	66	22	3	3	1		95
데이터 구축/컨설팅	8. 구축	1,003	398	82	75	26	10	1,594
	9. 컨설팅	824	140	17	18	3	2	1,004
데이터 서비스	10. 데이터 거래	134	60	14	9	7	1	225
	11. 정보 제공	2,044	385	82	77	21	8	2,617
	12. 데이터 분석 제공	54	21	2	4	1	3	85
합계		5,242	1,293	270	228	71	29	7,133

2) 표본설계

가. 표본크기 결정 및 표본 배분

- 최종 표본의 크기는 목표상대표준오차가 각 분류 내 종사자 규모 층에서 20% 이내가 되도록 총 1,000개의 사업체를 표본으로 결정하였다.

[그림1-1] 네이만 배분법에 의한 표본크기 결정 공식

$$n = \frac{\left(\sum_{h=1}^L W_h S_h \right)^2}{\sum_{h=1}^L W_h S_h^2 / N + \left(\alpha \sum_{h=1}^L W_h \bar{Y}_h \right)^2}$$

L : 층의 개수 (업종×규모)

S_h^2 : h 층의 종사자수 분산

W_h : 층별 가중치

N : 총 사업체수

$\bar{Y}_h$: h 층의 종사자수의 평균

$\alpha = \frac{\sqrt{\text{Var}(\bar{Y}_{st})}}{\bar{Y}_{st}}$: 목표상대표준오차

- 표본의 크기를 결정하기 위해 Neyman의 최적할당을 응용하였으며, 목표상대표준오차

$\alpha = \frac{\sqrt{\text{Var}(\bar{Y}_{st})}}{\bar{Y}_{st}} * 100(\%)$ 가 0.5% 이내가 되도록 총 표본의 크기를 1,000개로 결정하였다.

그리고 본 조사에서는 비용, 시간, 추정치의 상대표준오차 등 여러 가지 변수를 고려한 후 최종 표본의 크기를 목표상대표준오차가 1%인 경우로 결정하였다.⁸⁾

나. 표본의 할당

- 종사자규모 1,000인 이상인 29개 사업체는 전수층으로 처리하였다.
- 총 표본의 크기 1,000개 중에서 전수층의 29개를 제외하고 971개 사업체를 역등할당 방법을 이용하여 할당하였다.
- 할당방법별 상대표준오차를 계산하면, 아래 표에서 보는 바와 같이 역수($p=0.4$)인 경우가 업종별

8) 상대표준오차 1%는 추정치의 정도가 매우 높은 것으로 알려져 있음

종사자규모별 셀에서의 상대표준오차가 가장 적은 것으로 나타나므로 최종 표본의 할당으로 결정하였다.

[표1-10] 할당방법별 표본의 크기 및 상대표준오차

분류	종사자 규모	모집단 크기	Nyman		p=0.4		p=0.5		p=0.6	
			n	rse	n ⁹⁾	rse ¹⁰⁾	n	rse	n	rse
1.데이터 수집	1~9	274	12	29.1	21	22.0	19	23.2	18	23.8
	10~49	47	12	15.6	21	11.8	20	12.1	18	12.7
	50~99	13	3	8.8	13	4.2	11	4.6	9	5.1
	100~299	5	5	-	5	-	5	-	5	-
	300~999	2	2	-	2	-	2	-	2	-
	1000이상	-	-	-	-	-	-	-	-	-
2.데이터 설계	1~9	171	6	37.8	16	23.1	13	25.7	11	27.9
	10~49	36	10	16.4	20	11.6	18	12.3	16	13.0
	50~99	10	4	11.4	10	-	10	-	9	7.6
	100~299	4	4	-	4	-	4	-	4	-
	300~999	2	2	-	2	-	2	-	2	-
	1000이상	1	1	-	1	-	1	-	1	-
3.DBMS	1~9	109	5	44.8	15	25.9	13	27.8	11	30.2
	10~49	14	3	23.3	12	11.7	10	12.8	8	14.3
	50~99	4	2	22.5	4	-	4	-	4	-
	100~299	3	0	-	3	-	3	-	2	1.8
	300~999	4	4	-	4	-	4	-	4	-
	1000이상	1	1	-	1	-	1	-	1	-
4.데이터 관리	1~9	302	16	24.3	24	19.9	23	20.3	22	20.8
	10~49	115	30	9.1	30	9.1	31	9.0	31	9.0
	50~99	30	10	6.4	19	4.7	17	4.9	16	5.1
	100~299	21	19	5.8	21	-	21	-	21	-
	300~999	2	2	-	2	-	2	-	2	-
	1000이상	2	2	-	2	-	2	-	2	-
5.데이터 품질	1~9	136	5	45.1	15	26.0	13	28.0	11	30.4
	10~49	15	4	22.0	13	12.2	11	13.3	9	14.7
	50~99	5	2	16.1	5	-	5	-	5	-
	100~299	2	2	-	2	-	2	-	2	-
	300~999	-	-	-	-	-	-	-	-	-
	1000이상	-	-	-	-	-	-	-	-	-

9) 표본의 크기

10) 상대표준오차(relative standard error)

(계속)

분류	종사자 규모	모집단 크기	Nyman		p=0.4		p=0.5		p=0.6	
			n	rse	n	rse	n	rse	n	rse
6.데이터 분석	1~9	125	6	40.4	16	24.7	14	26.4	12	28.5
	10~49	40	11	15.5	20	11.5	19	11.8	17	12.5
	50~99	8	3	14.7	8	-	8	-	8	-
	100~299	7	7	-	7	-	7	-	7	-
	300~999	2	2	-	2	-	2	-	2	-
	1000이상	1	1	-	1	-	1	-	1	-
7.데이터 플랫폼	1~9	66	4	42.2	14	22.5	12	24.3	9	28.1
	10~49	22	5	21.7	15	12.5	13	13.4	11	14.6
	50~99	3	0	-	3	-	3	-	3	-
	100~299	3	3	-	3	-	3	-	3	-
	300~999	1	1	-	1	-	1	-	1	-
	1000이상	-	-	-	-	-	-	-	-	-
8.구축	1~9	1003	55	13.4	39	15.9	42	15.3	45	14.8
	10~49	398	99	4.6	49	6.5	56	6.1	64	5.7
	50~99	82	29	3.9	30	3.9	30	3.9	31	3.8
	100~299	75	75	-	48	4.6	55	4.3	62	4.0
	300~999	26	26	-	26	-	26	-	26	-
	1000이상	10	10	-	10	-	10	-	10	-
9.컨설팅	1~9	824	39	15.9	34	17.0	35	16.7	36	16.5
	10~49	140	35	8.5	33	8.8	34	8.6	34	8.6
	50~99	17	7	8.6	17	5.5	15	5.9	13	6.3
	100~299	18	18	-	18	-	18	-	18	-
	300~999	3	3	-	3	-	3	-	3	-
	1000이상	2	2	-	2	-	2	-	2	-
10.데이터 거래	1~9	134	8	33.4	18	22.3	16	23.6	14	25.2
	10~49	60	16	12.3	23	10.3	22	10.5	21	10.8
	50~99	14	3	8.8	12	4.4	10	4.8	8	5.4
	100~299	9	9	-	9	-	9	-	9	-
	300~999	7	7	-	7	-	7	-	7	-
	1000이상	1	1	-	1	-	1	-	1	-
11.정보 제공	1~9	2044	101	8.9	50	12.7	57	11.9	64	11.2
	10~49	385	94	5.3	48	7.4	55	6.9	62	6.5
	50~99	82	31	4.1	31	4.1	31	4.1	31	4.1
	100~299	77	77	-	48	4.4	55	4.2	62	3.9
	300~999	21	21	-	21	-	21	-	21	-
	1000이상	8	8	-	8	-	8	-	8	-
12.데이터 분석 제공	1~9	54	3	44.9	13	21.6	10	24.6	8	27.5
	10~49	21	5	19.3	15	11.1	12	12.4	10	13.6
	50~99	2	0	-	2	-	2	-	2	-
	100~299	4	3	8.4	4	-	4	-	4	-
	300~999	1	1	-	1	-	1	-	1	-
	1000이상	3	3	-	3	-	3	-	3	-

[표1-11] 표본 분포

분류	종사자 규모						총합
	1-9인	10-49인	50-99인	100-299인	300-999인	1,000인 이상	
1. 데이터 수집	21	21	13	5	2	-	62
2. 데이터 설계	16	20	10	4	2	1	53
3. DBMS	15	12	4	3	4	1	39
4. 데이터 관리	24	30	19	21	2	2	98
5. 데이터 품질	15	13	5	2	-	-	35
6. 데이터 분석	16	20	8	7	2	1	54
7. 데이터 플랫폼	14	15	3	3	1	-	36
8. 구축	39	49	30	48	26	10	202
9. 컨설팅	34	33	17	18	3	2	107
10. 데이터 거래	18	23	12	9	7	1	70
11. 정보 제공	50	48	31	48	21	8	206
12. 데이터 분석 제공	13	15	2	4	1	3	38
합계	275	299	154	172	71	29	1,000

1.4 모수 추정

1) 회수현황 및 이상치 처리

가. 회수현황

- 2017년 데이터산업 현황조사를 실시한 결과, 데이터솔루션, 데이터구축, 데이터컨설팅, 데이터서비스 4개 분야의 표본할당 사업체 수는 1,000개이며, 이 중 응답한 사업체 수는 1,004개로 나타났다. 데이터산업 세부 산업분야별로 보면, 데이터솔루션이 274개, 데이터구축 262개, 데이터컨설팅 86개, 데이터서비스는 382개를 조사완료 하였다.

[표1-12] 조사기업 분포

구분	종사자 규모						총합
	1-9인	10-49인	50-99인	100-299인	300-999인	1,000인 이상	
데이터솔루션	82	127	34	24	6	1	274
데이터구축	56	137	32	25	12	-	262
데이터컨설팅	29	37	7	12	-	1	86
데이터서비스	145	155	40	28	12	2	382
합계	312	456	113	89	30	4	1004

- 사업 분야별로 목표 표본 대비 회수율은 데이터솔루션과 컨설팅을 제외하고 100% 이상의 회수율을 달성하였다.

[표1-13] 데이터산업 현황조사 회수율

구분	종사자 규모						총합
	1-9인	10-49인	50-99인	100-299인	300-999인	1,000인 이상	
데이터솔루션	67.8%	96.9%	54.8%	53.3%	46.2%	20.0%	72.7%
데이터구축	143.6%	279.6%	106.7%	52.1%	46.2%	-	129.7%
데이터컨설팅	85.3%	112.1%	41.2%	66.7%	-	50.0%	80.4%
데이터서비스	179.0%	180.2%	88.9%	45.9%	41.4%	16.7%	121.7%
합계	113.5%	152.5%	73.4%	51.7%	42.3%	13.8%	100.4%

나. 항목 무응답 처리

- 조사표 점검 시 조사가 되어야 하는 항목(매출액, 종사자 수) 중 일부가 누락된 경우 그 이유를 반드시 확인하여 누락된 사유를 작성하였다.
- 실사관리자는 조사표를 매일 점검하며, 주요항목이 누락된 경우에는 해당사업체에 직접 전화를 걸어 사유별 적절한 대응을 하였다. 실사관리자가 2~3회 이상 반복조사를 실시하여도 응답을 거부한 경우에는 이를 취합하여 1주일 단위로 조사담당자에게 보고하여 조치를 취하였다.

다. 항목 간 논리 점검

- 조사표 검토 시 항목 간에 내용이 연계되는 경우가 있으며, 이를 1차적으로 실사관리자가 현장에서 검토하고, 2차적으로 조사담당자가 검토하였다.

[표1-14] 항목 간 논리 점검 예시

구분	기술등급별 현재 인력			전체 데이터직무 인력	인력비중
	①초급	②중급	③고급		
DA	2	5	5	12	41.4%
데이터 개발자	3	4	5	12	41.4%
데이터 엔지니어	0	3	0	3	10.3%
데이터 분석가	0	1	1	2	6.9%
합계	5	13	11	29	100.0%

라. 이상치 파악 및 처리

- 조사항목 매출액, 종사자 수에 대해서는 이상치를 파악하여 처리하였다.
- 세부분야 매출액이 종사자 수에 비해 지나치게 적거나 많은 경우 이상치 처리 기준에 의거하여 확인하였고, 소분류별로 1인당 평균 매출액이 평균보다 2배 이상 큰 경우, 전년도 자료 확인 및 재조사를 실시하였다.
 - 이상치 기준 : $(\text{평균} \pm 2 \times \text{표준편차})$ 의 범위를 벗어나는 값

2) 단위 무응답 추정

- 추정 대상이 되는 조사 사업체 수의 매출액 및 종사자 무응답 현황은 다음과 같다. 2017년 매출액이 조사되지 않는 사업체는 26개이며, 종사자 수가 조사되지 않는 사업체는 15개로 나타났다.

[표1-15] 추정 대상 무응답 사업체 수

구분	매출액 무응답 사업체		종사자 수 무응답 사업체
	2016년	2017년(예상)	
데이터솔루션	6	6	5
데이터구축	3	3	3
데이터컨설팅	4	4	2
데이터서비스	13	13	5
합계	26	26	15

- 주요 항목에 대해서는 중소기업현황정보시스템¹¹⁾, 전자공시시스템(DART)¹²⁾ 및 나이스평가정보(주)의 Rm¹³⁾을 조회하여 최신 정보를 우선적으로 수집하여 인용하였다.

가. 매출액 추정

- 매출액 무응답으로 처리된 사업체는 대부분 50인 미만 사업체이며, 종사자 수가 동일한 사업체가 5개 이상인 경우 이상치를 제외하고 평균치로 대체하여 추정하였다.

나. 종사자 수 추정

- 종사자 수 무응답 사업체는 2016년 자료, 전자공시시스템(DART), 나이스평가정보(주)의 Rm1 자료를 활용하여 대체하였다.

11) <http://sminfo.smba.go.kr>

12) <http://dart.fss.or.kr>

13) <http://www.rm1.co.kr>

3) 최종 추정치

○ 사업 분야별 총계 추정치

$$\hat{\tau}_{\text{업종}} = {}_c\tau_{1000} + \sum_{h=1}^{12} {}_s\hat{\tau}_h = \sum_{i=1}{{}_cN_{1000}} Y_{1000,i} + \sum_{h=1}^{12} \left(\frac{{}_sN_h}{{}_sn_h} \right) \sum_{i=1}{{}_cn_h} y_{hi}$$

여기에서 $\hat{\tau}_{\text{업종}}$: 사업분야의 총계추정치

${}_c\tau_{1000}$: 전수층 총계

${}_s\hat{\tau}_h$: 종사자규모 h의 표본층 총계추정치

$Y_{1000,i}$: 전수층에 속하는 사업체의 특성

${}_cN_{1000}$: 전수층의 크기

${}_sN_h$: 종사자규모 h의 표본층의 크기

${}_sn_h$: 종사자규모 h의 표본층 표본크기

y_{hi} : 종사자규모 h의 표본층 표본사업체의 특성

○ 분산 추정치

$$\hat{V}(\hat{\tau}_{\text{업종}}) = \sum_{h=1}^{12} {}_sN_h^2 \left(\frac{{}_sN_h - {}_sn_h}{{}_sN_h} \right) \frac{{}_s^2}{{}_sn_h}$$

여기에서 ${}_s^2 = \frac{1}{{}_sn_h - 1} \sum_{i=1}{{}_sn_h} (y_{hi} - \bar{y}_h)^2$ 은 표본층의 분산추정치

○ 허용오차 (95% 신뢰구간)

$$B = 1.96 * \sqrt{\hat{V}(\hat{\tau}_{\text{업종}})}$$

[표1-16] 데이터산업 매출액 상대표준오차, 신뢰수준

대분류	중분류	2016년 매출액 (억 원)	표준오차 (억 원)	상대표준오차 (%)	신뢰수준 (%)
데이터솔루션	데이터 수집	1,345	0.2	0.5	95
	데이터 설계	225	0.1	0.9	95
	DBMS	6,148	3.1	2.6	95
	데이터 관리	4,022	0.3	0.8	95
	데이터 품질	1,120	1.4	2.2	95
	데이터 분석	1,249	0.9	2.1	95
	데이터 플랫폼	1,611	3.1	4.0	95
데이터 구축/컨설팅	데이터구축	54,571	1.0	0.5	95
	데이터컨설팅	1,279	0.0	0.3	95
데이터서비스	데이터 거래	2,594	0.6	1.2	95
	정보 제공	59,854	1.3	0.6	95
	데이터 분석 제공	3,529	9.5	5.9	95

[표1-17] 데이터산업 직접매출액 상대표준오차, 신뢰수준

대분류	중분류	2016년 매출액 (억 원)	표준오차 (억 원)	상대표준오차 (%)	신뢰수준 (%)
데이터솔루션	데이터 수집	1,345	0.2	0.5	95
	데이터 설계	225	0.1	0.9	95
	DBMS	6,148	3.1	2.6	95
	데이터 관리	4,022	0.3	0.8	95
	데이터 품질	1,120	1.4	2.2	95
	데이터 분석	1,249	0.9	2.1	95
	데이터 플랫폼	1,611	3.1	4.0	95
데이터 구축/컨설팅	데이터구축	26,596	0.7	0.7	95
	데이터컨설팅	1,279	0.0	0.3	95
데이터서비스	데이터 거래	2,354	0.5	1.1	95
	정보 제공	11,523	0.2	0.5	95
	데이터 분석 제공	3,051	9.4	6.8	95

[표1-18] 데이터직무 인력 상대표준오차, 신뢰수준

대분류		2017년 데이터직무인력 (명)	표준오차 (명)	상대표준오차 (%)	신뢰수준 (%)
데이터솔루션		10,291	0.1	0.4	95
데이터 구축/컨설팅	데이터구축	31,693	0.5	0.4	95
	데이터컨설팅	5,823	0.3	0.5	95
데이터서비스		29,298	0.2	0.2	95
일반산업 ¹⁴⁾		32,215	0.1	0.1	95

14) 일반산업의 데이터직무 인력현황 파악을 위해 한국정보화진흥원에서 실시한 '2017년 빅데이터 시장 현황조사' 데이터를 활용하여 추정

4) 데이터산업 현황조사 이용 시 유의사항

- 데이터솔루션, 데이터구축 및 컨설팅, 데이터서비스 산업에서 데이터 관련 사업 분야가 1개 이상인 사업체는 해당 사업체의 사업 분야 매출비중을 기준으로 사업체 수를 집계하였다.
- 데이터솔루션, 데이터구축 및 컨설팅, 데이터서비스 중분류 사업을 동시에 영위하는 사업체가 있어 중분류 상에 중복이 집계되었다. 매출액의 경우 각 영역의 데이터 관련 매출만 집계되기 때문에 중분류별 매출액에는 중복이 계산되지 않았다. 다만, 시장현황 및 정책수요 결과에 중복이 허용된 응답 수를 산출하여 작성하였다.
- 매년 조사 시점 매출액은 예상 매출액으로 산출하고, 차기년도에 매출액 확정치를 조사하여 반영하고 있다. 즉 2017년 보고서에는 2016년 확정치와 2017년 예상치가 수록되므로 2017년 보고서상의 2016년 시장 규모가 전년도 보고서상의 규모와 상이할 수 있음을 밝혀둔다.
- 2017년 보고서는 데이터와 직접적으로 관련이 있는 매출은 물론 DB시스템 구축, 데이터를 매개로 하는 광고매출 등 기존의 DB산업현황조사에서부터 유지되어 오던 DB 기반의 간접매출을 모두 포함하는 광의의 개념으로서 '데이터산업 시장 규모'와 데이터와 직접적으로 관련이 있는 매출 중심의 '데이터산업 직접매출 시장 규모'를 구분하여 수록하고 있다.
- 데이터솔루션, 데이터구축 및 컨설팅, 데이터서비스 산업의 시장 규모 분석은 중분류로 분석하였으며, 이전 년도와 비교가 가능하도록 하였다.
- 본 보고서의 '제5부 데이터 유통 현황'은 데이터서비스 기업을 대상으로 조사한 결과이다.
- 보고서에 수록된 수치는 소수점 미만을 반올림한 것으로 백분율의 합이 100이 안 될 수 있다.
- 보고서에 수록된 통계수치를 이용할 경우 주요항목별 표준오차를 확인해 참고할 것을 권장한다.
- 보고서에 사용한 부호는 다음과 같다.
 - 0 : 단위 미만, - : 해당숫자 없음

The background features a stylized city skyline in shades of blue and grey on the left side. The rest of the background is composed of large, overlapping geometric shapes in light blue, white, and grey, creating a modern, abstract design.

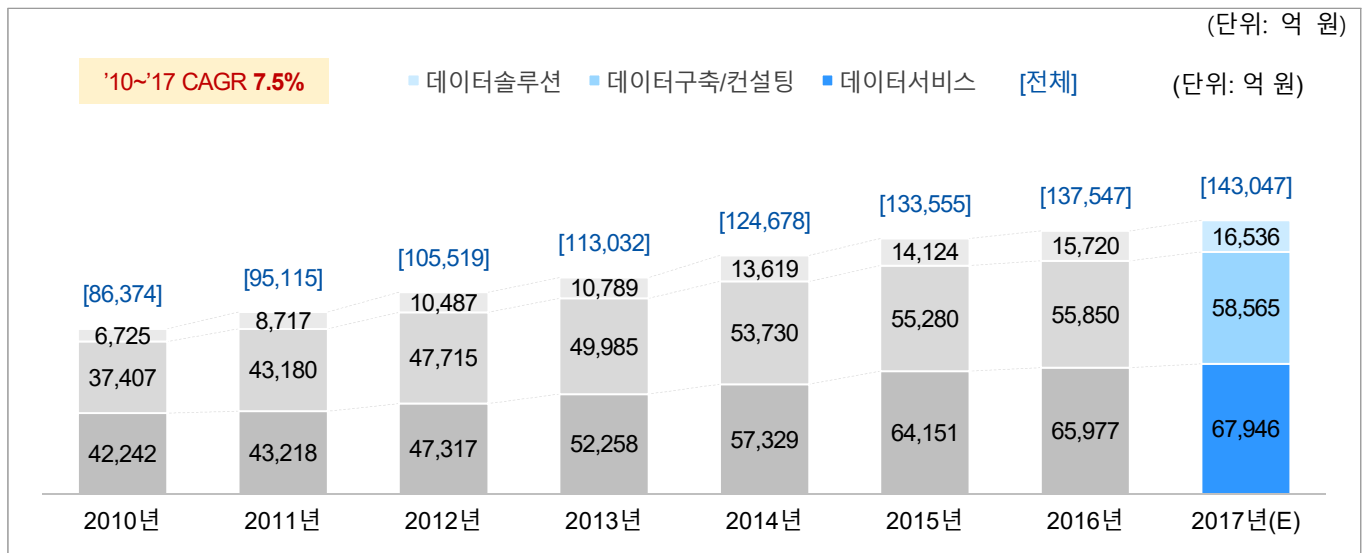
제2부. 데이터산업 시장 규모

2.1 데이터산업 시장 규모

1) 전체 시장

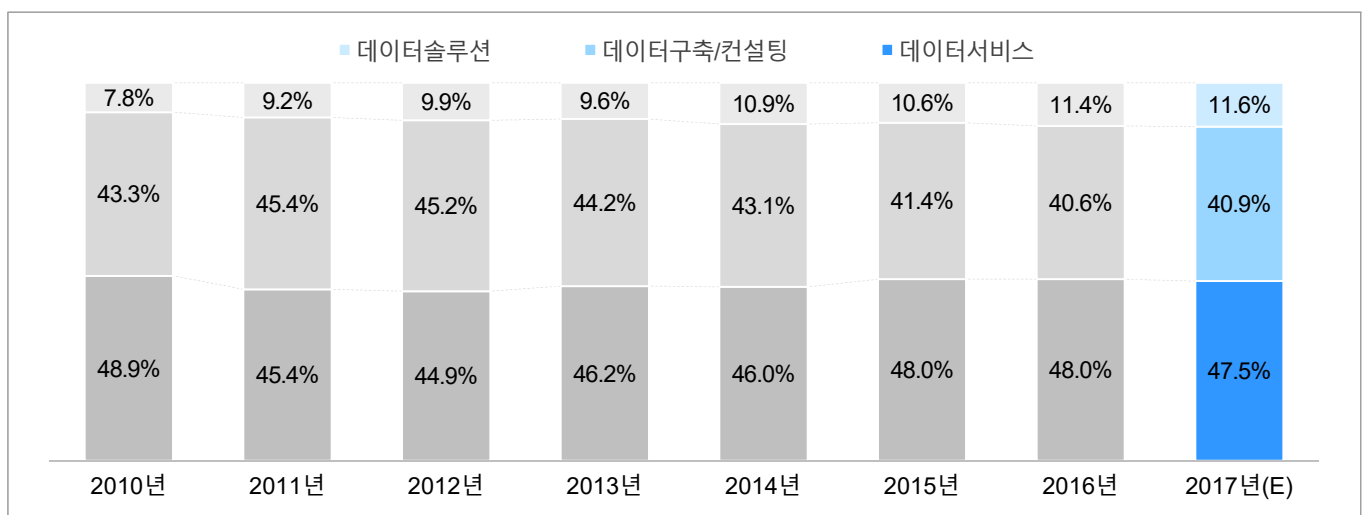
- 국내 데이터산업 시장 규모를 분석한 결과¹⁵⁾, 2016년은 13조 7,547억 원이며, 2017년에는 14조 3,047억 원으로 전년 대비 4.0% 성장할 것으로 나타났다. 2010년 이후 연평균증가율은 7.5%로 매년 성장세를 유지하고 있는 것으로 나타났다.

[그림2-1] 데이터산업 시장 규모



- 2017년 기준, 부문별 규모를 보면, 데이터서비스 시장이 47.5%로 가장 큰 비중을 차지하였고, 이어서 데이터구축/컨설팅 시장 40.9%, 데이터솔루션 11.6% 순으로 나타났다.

[그림2-2] 데이터산업 부문별 시장 규모 비중



15) 데이터산업 시장은 전통적인 DB산업을 아우르는 기존 조사 범위를 적용한 것으로, 데이터를 매개로하는 광고 매출, DB시스템 구축 용역 매출 등 DB 관련 간접 매출까지 포함하여 산출한 광의의 데이터산업 시장 규모임

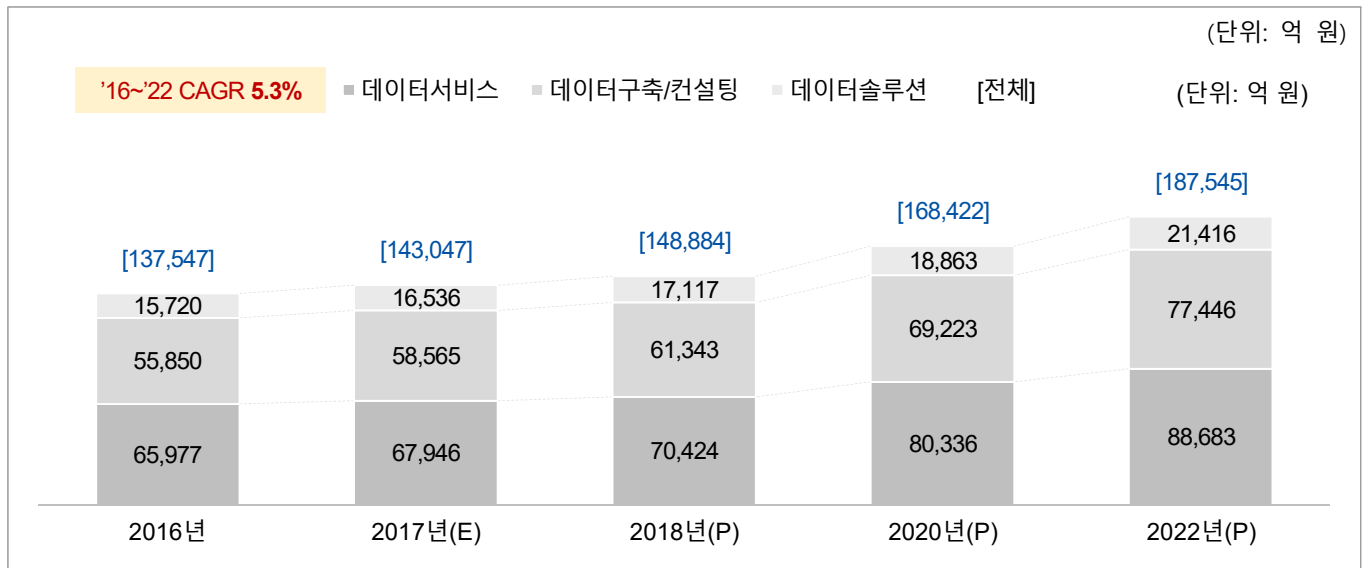
[표2-1] 데이터산업 시장 규모 총괄

구분	2010년	2011년	2012년	2013년	2014년	2015년	2016년	2017년 (E)	증감률 '16~'17	CAGR '10~'17
데이터솔루션	6,725	8,717	10,487	10,789	13,619	14,124	15,720	16,536	5.2%	13.7%
데이터구축/컨설팅	37,407	43,180	47,715	49,985	53,730	55,280	55,850	58,565	4.9%	6.6%
데이터서비스	42,242	43,218	47,317	52,258	57,329	64,151	65,977	67,946	3.0%	7.0%
전체	86,374	95,115	105,519	113,032	124,678	133,555	137,547	143,047	4.0%	7.5%

(단위: 억 원)

- 향후 데이터산업은 2022년까지 연평균증가율 5.3%의 성장세를 유지하며 18조 원대 시장 진입을 할 것으로 전망된다.

[그림2-3] 데이터산업 시장 전망



[표2-2] 데이터산업 시장 전망

구분	2016년	2017년(E)	2018년(P)	2019년(P)	2020년(P)	2021년(P)	2022년(P)	CAGR '16~'22
데이터솔루션	15,720	16,536	17,117	18,010	18,863	20,234	21,416	5.3%
데이터구축/컨설팅	55,850	58,565	61,343	65,391	69,223	73,052	77,446	5.6%
데이터서비스	65,977	67,946	70,424	75,869	80,336	83,832	88,683	5.1%
전체	137,547	143,047	148,884	159,270	168,422	177,118	187,545	5.3%

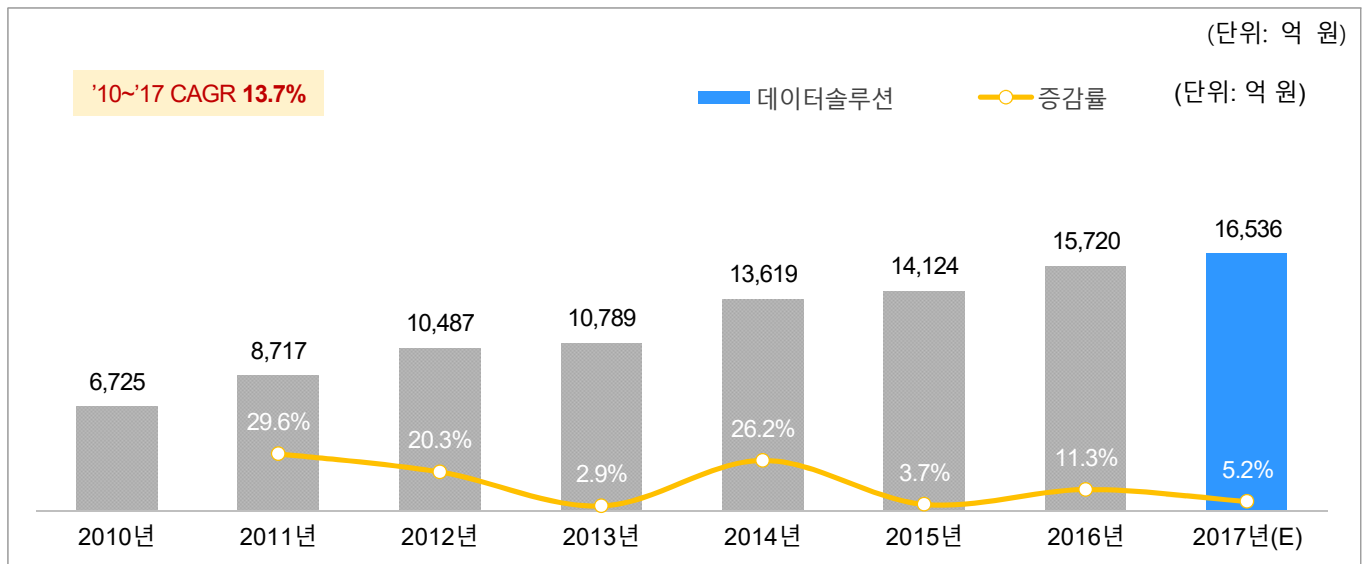
(단위: 억 원)

2) 부문별 시장

가. 데이터솔루션 시장

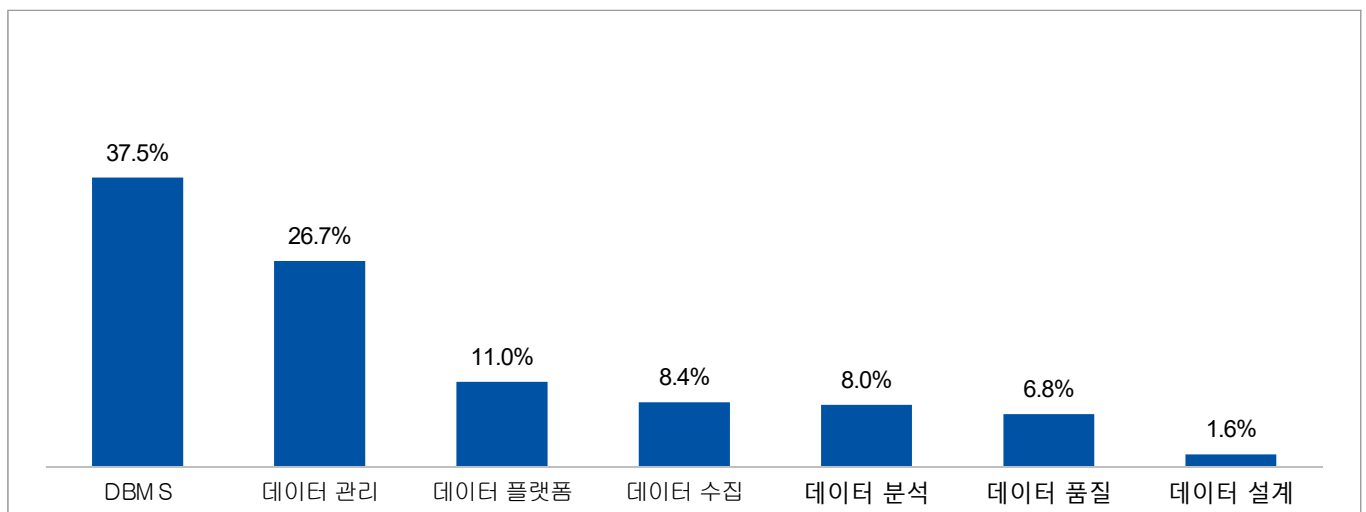
- 데이터솔루션 시장 규모는 2016년 1조 5,720억 원이며, 2017년에는 1조 6,536억 원으로 전년 대비 5.2% 성장할 것으로 나타났다.
- 2017년에는 전년도 대비 공공 부분의 비중은 줄어든 반면, 통신/미디어 분야의 투자 비중이 높아진 것으로 나타났다.

[그림2-4] 데이터솔루션 시장 규모



- 2017년 데이터솔루션 중분류 시장에서 DBMS 분야가 37.5%로 가장 큰 비중을 차지하고 있으며, 다음으로 데이터 관리, 데이터 플랫폼, 데이터 수집, 데이터 분석, 데이터 품질의 순으로 나타났다.

[그림2-5] 2017년 데이터솔루션 중분류별 시장 규모 비중



[표2-3] 데이터솔루션 중분류별 시장 규모

구분	2010년	2011년	2012년	2013년	2014년	2015년	2016년	2017년 (E)	증감률 '16~'17	CAGR '10~'17
데이터 수집	448	570	617	622	1,076	1,115	1,345	1,382	2.8%	17.5%
데이터 설계	129	136	151	166	202	207	225	269	19.6%	11.1%
DBMS	3,169	4,272	5,301	5,488	5,502	5,727	6,148	6,192	0.7%	10.0%
데이터 관리	1,237	1,516	2,075	2,489	3,446	3,574	4,022	4,417	9.8%	19.9%
데이터 품질	686	949	942	855	883	918	1,120	1,127	0.6%	7.3%
데이터 분석	1,056	1,274	1,401	1,169	1,121	1,157	1,249	1,328	6.3%	3.3%
데이터 플랫폼	-	-	-	-	1,389	1,426	1,611	1,821	13.0%	9.4% ¹⁶⁾
데이터솔루션 전체	6,725	8,717	10,487	10,789	13,619	14,124	15,720	16,536	5.2%	13.7%

(단위: 억 원)

○ 라이선스, 개발, 유지보수 영역별 규모는 다음과 같다.

[표2-4] 데이터솔루션 영역별 시장 규모

구분		2010년	2011년	2012년	2013년	2014년	2015년	2016년	2017년 (E)	증감률 '16~'17	CAGR '10~'17
데이터 수집	라이선스	370	433	477	492	316	225	277	289	4.3%	14.6%
	개발					507	523	681	673	-1.2%	
	유지보수	78	137	140	130	253	367	387	420	8.5%	
데이터 설계	라이선스	111	117	130	144	25	33	42	50	19.0%	8.1%
	개발					141	133	125	142	13.6%	
	유지보수	18	19	21	22	36	41	58	77	32.8%	
DBMS	라이선스	2,531	3,590	4,718	4,722	2,698	2,712	2,880	2,948	2.4%	10.4%
	개발					1,987	2,101	2,184	2,127	-2.6%	
	유지보수	638	682	583	766	817	914	1,084	1,117	3.0%	
데이터 관리	라이선스	953	1,212	1,692	1,987	1,395	363	452	474	4.9%	14.5%
	개발					1,189	1,796	1,902	1,987	4.5%	
	유지보수	281	304	383	502	862	1,415	1,668	1,956	17.3%	
데이터 품질	라이선스	481	662	630	587	436	131	179	210	17.3%	4.2%
	개발					243	354	438	433	-1.1%	
	유지보수	205	287	312	268	204	433	503	484	-3.8%	
데이터 분석	라이선스	841	1,111	1,319	967	104	116	190	207	8.9%	3.1%
	개발					821	846	830	836	0.7%	
	유지보수	215	163	82	202	196	195	229	285	24.5%	
데이터 플랫폼 ¹⁷⁾	라이선스	-	-	-	-	391	361	352	432	22.7%	3.4%
	개발	-	-	-	-	746	823	958	1,068	11.5%	
	유지보수	-	-	-	-	252	242	301	321	6.6%	
데이터 솔루션 전체	라이선스	5,287	7,125	8,966	8,899	5,365	3,941	4,372	4,610	5.4%	12.3%
	개발					5,634	6,576	7,118	7,266	2.1%	
	유지보수	1,435	1,592	1,521	1,890	2,620	3,607	4,230	4,660	10.2%	

(단위: 억 원)

16) 2014년~2017년 CAGR

17) 2014년~2017년 CAGR

- 업종별 매출 비중을 살펴보면, 공공 분야와 유통/서비스가 가장 큰 수요층이고, 전년 대비 통신/미디어의 비중이 높아진 것으로 나타났다.

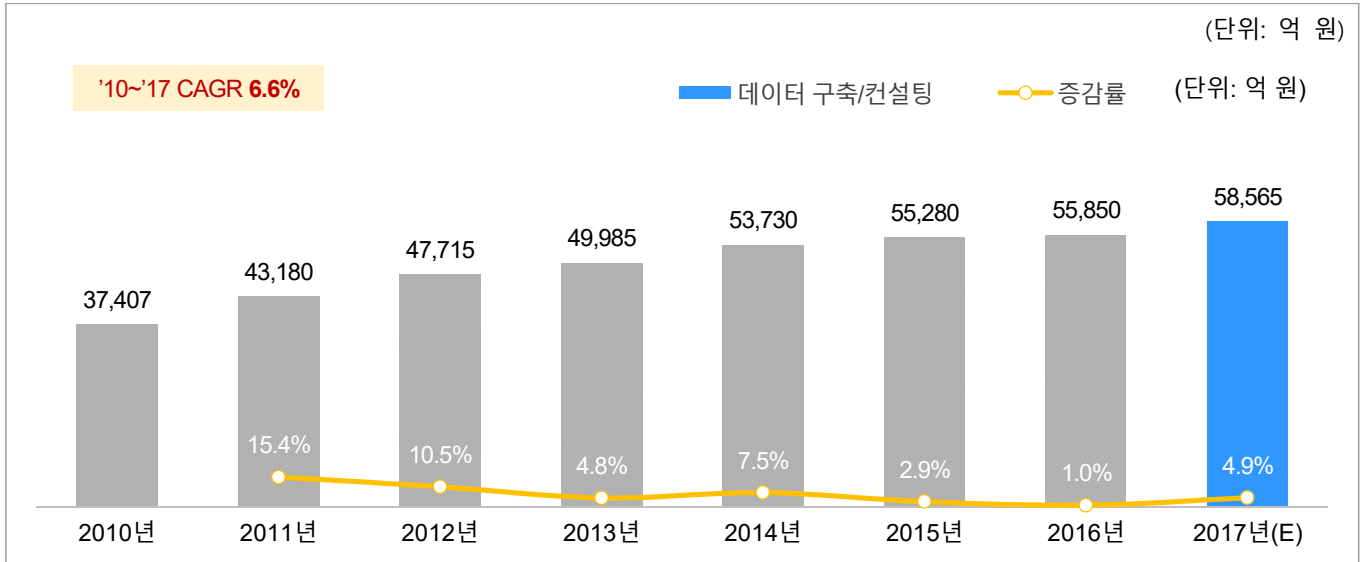
[표2-5] 데이터솔루션 업종별 매출 비중

구분		공공	금융	제조/건설	유통/ 서비스	통신/ 미디어	의료	전체
데이터 수집	2016년	43.7%	10.4%	14.8%	27.5%	2.8%	0.8%	100.0%
	2017년(E)	21.3%	11.7%	23.0%	23.9%	17.8%	2.3%	100.0%
데이터 설계	2016년	25.8%	12.0%	20.5%	31.5%	6.9%	3.3%	100.0%
	2017년(E)	28.2%	11.1%	17.2%	26.9%	13.9%	2.7%	100.0%
DBMS	2016년	49.6%	7.5%	16.3%	17.7%	5.0%	3.9%	100.0%
	2017년(E)	39.3%	5.0%	12.7%	26.1%	16.9%	0.0%	100.0%
데이터 관리	2016년	42.3%	7.7%	17.2%	24.3%	4.4%	4.1%	100.0%
	2017년(E)	31.1%	12.1%	14.5%	23.6%	14.7%	4.0%	100.0%
데이터 품질	2016년	22.4%	29.4%	22.8%	14.1%	9.6%	1.7%	100.0%
	2017년(E)	7.5%	6.6%	33.3%	26.8%	24.6%	1.2%	100.0%
데이터 분석	2016년	41.3%	12.5%	17.2%	17.5%	8.2%	3.3%	100.0%
	2017년(E)	23.0%	7.3%	28.6%	21.4%	9.1%	10.6%	100.0%
데이터 플랫폼	2016년	48.1%	10.8%	8.9%	16.8%	13.9%	1.5%	100.0%
	2017년(E)	21.3%	11.0%	16.5%	30.2%	20.0%	1.0%	100.0%
데이터 솔루션 전체	2016년	39.0%	12.9%	16.8%	21.4%	7.2%	2.7%	100.0%
	2017년(E)	26.7%	10.4%	18.9%	24.7%	15.7%	3.6%	100.0%

나. 데이터구축/컨설팅 시장

- 데이터구축/컨설팅 시장 규모는 2016년 5조 5,850억 원이며, 2017년은 5조 8,565억 원으로 전년 대비 4.9% 상승할 것으로 나타났으며, 2010년 이후 연평균증가율은 6.6%이다.

[그림2-6] 데이터구축/컨설팅 시장 규모



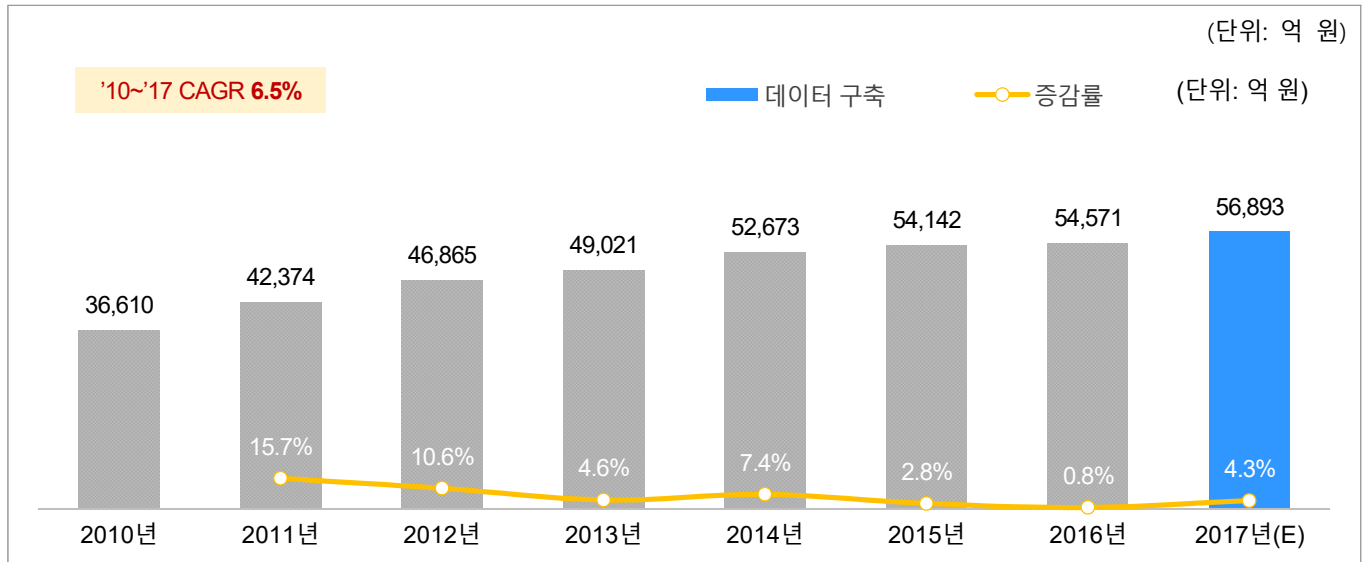
[표2-6] 데이터 구축/컨설팅 시장 규모

구분	2010년	2011년	2012년	2013년	2014년	2015년	2016년	2017년 (E)	증감률 '16~'17	CAGR '10~'17
데이터 구축	36,610	42,374	46,865	49,021	52,673	54,142	54,571	56,893	4.3%	6.5%
데이터 컨설팅	797	806	850	964	1,057	1,138	1,279	1,672	30.7%	11.2%
데이터구축/컨설팅 전체	37,407	43,180	47,715	49,985	53,730	55,280	55,850	58,565	4.9%	6.6%

(단위: 억 원)

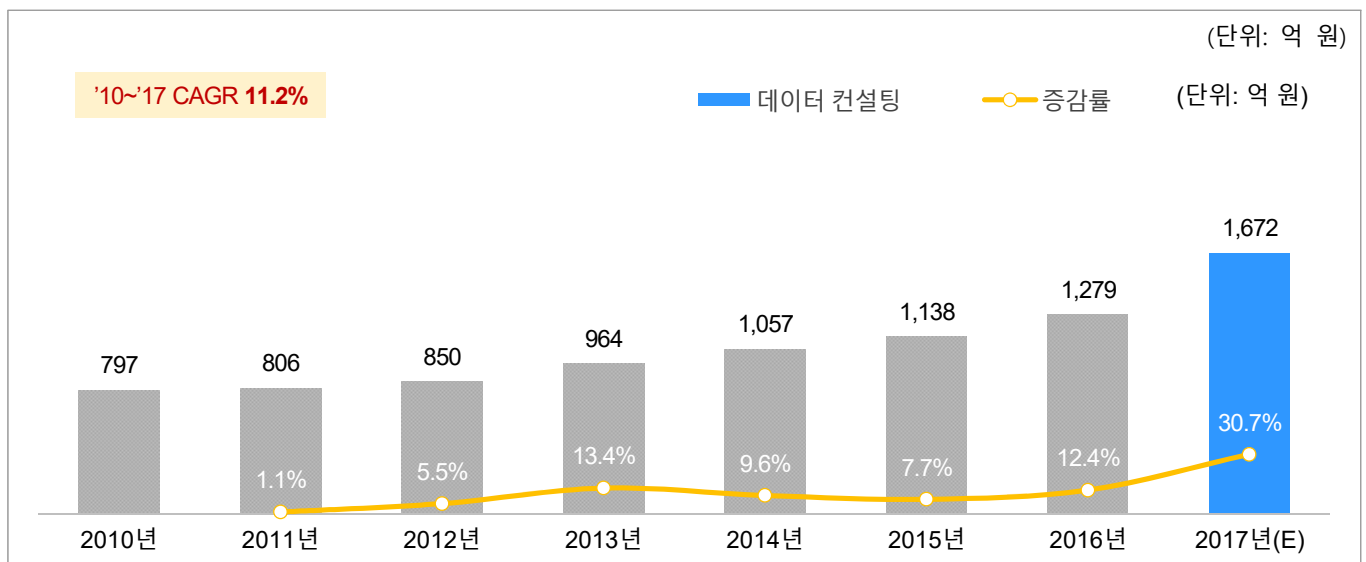
- 데이터구축 시장은 2016년 5조 4,571억 원이며, 2017년에는 5조 6,893억 원으로 전년 대비 4.3% 상승할 것으로 나타났다. 2010년 이래 연평균성장률은 6.5%이다.

[그림2-7] 데이터구축 시장 규모



- 데이터컨설팅 시장은 2016년 1,279억 원이며, 2017년에는 1,672억 원으로 2016년 대비 30.7% 성장할 것으로 조사되었다. 2010년부터 연평균성장률 11.2%로 성장하고 있다.

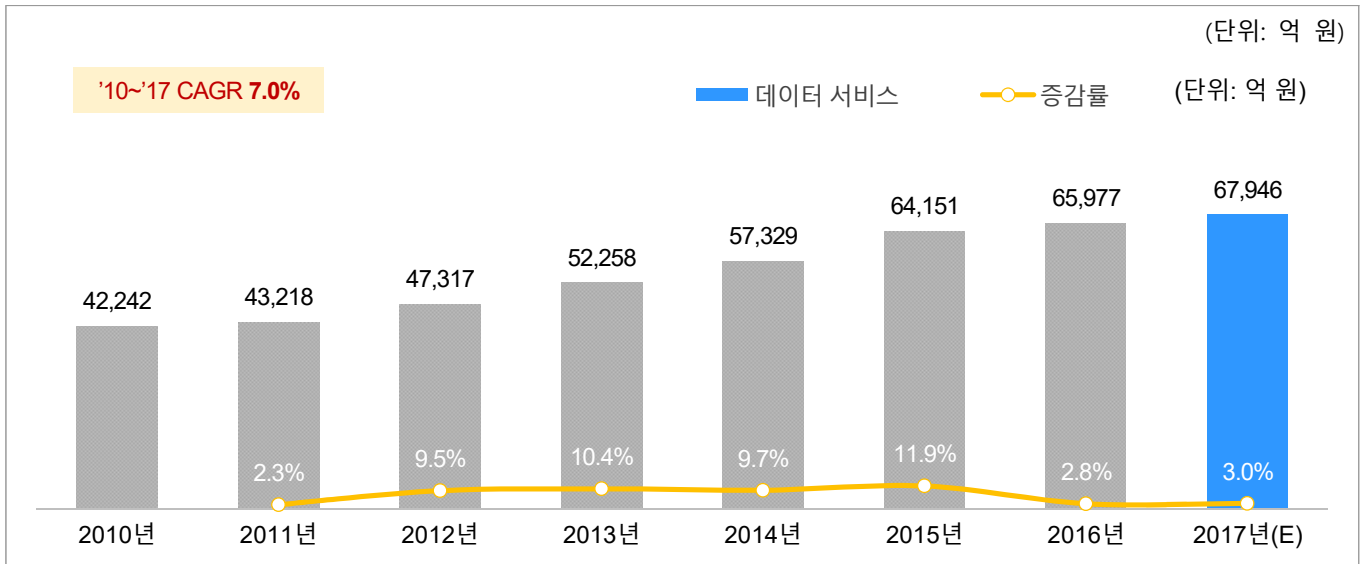
[그림2-8] 데이터컨설팅 시장 규모



다. 데이터서비스 시장

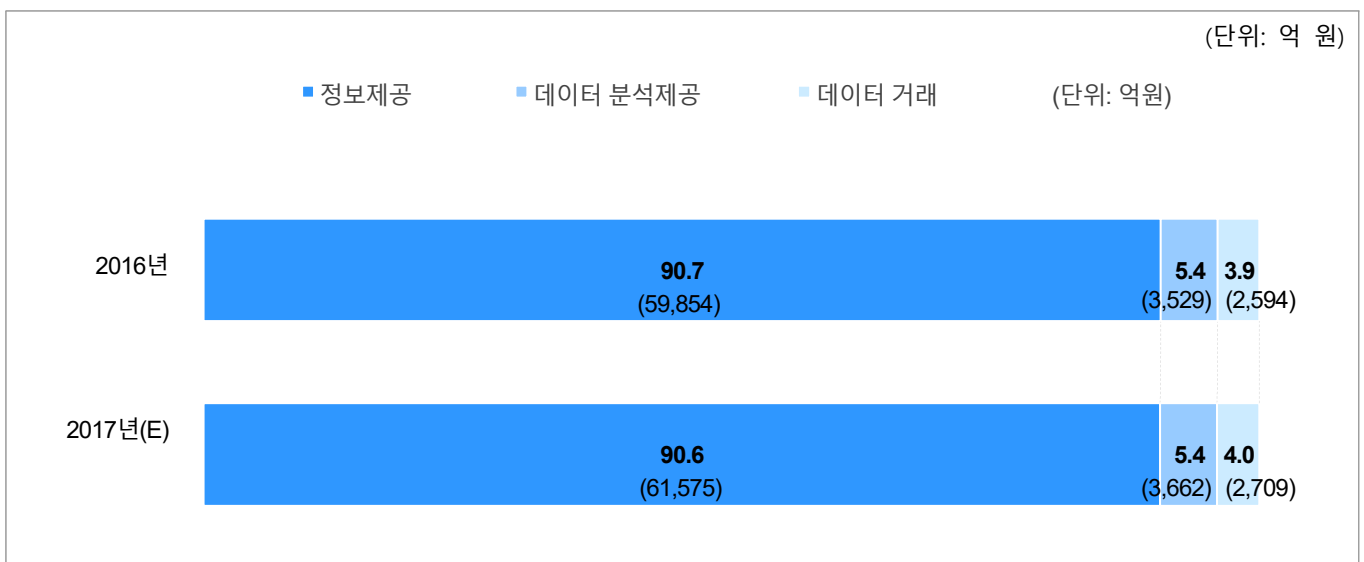
- 데이터서비스 시장은 2016년 6조 5,977억 원이며, 2017년에는 6조 7,946억 원으로 전년 대비 3.0% 증가할 것으로 조사되었다. 2010년 이래 연평균성장률 7.0%로 성장하고 있다.

[그림2-9] 데이터서비스 시장 규모



- 데이터서비스 중분류별 시장 규모를 보면, 정보 제공 시장 규모가 90% 가량을 차지하고 있다.
- 데이터 거래 시장은 2016년 2,594억 원이며, 2017년은 전년 대비 4.4% 증가한 2,709억 원으로 나타났다, 데이터 분석 제공 시장은 2017년 3,662억 원으로 나타났다.

[그림2-10] 데이터서비스 중분류별 시장 규모 비중



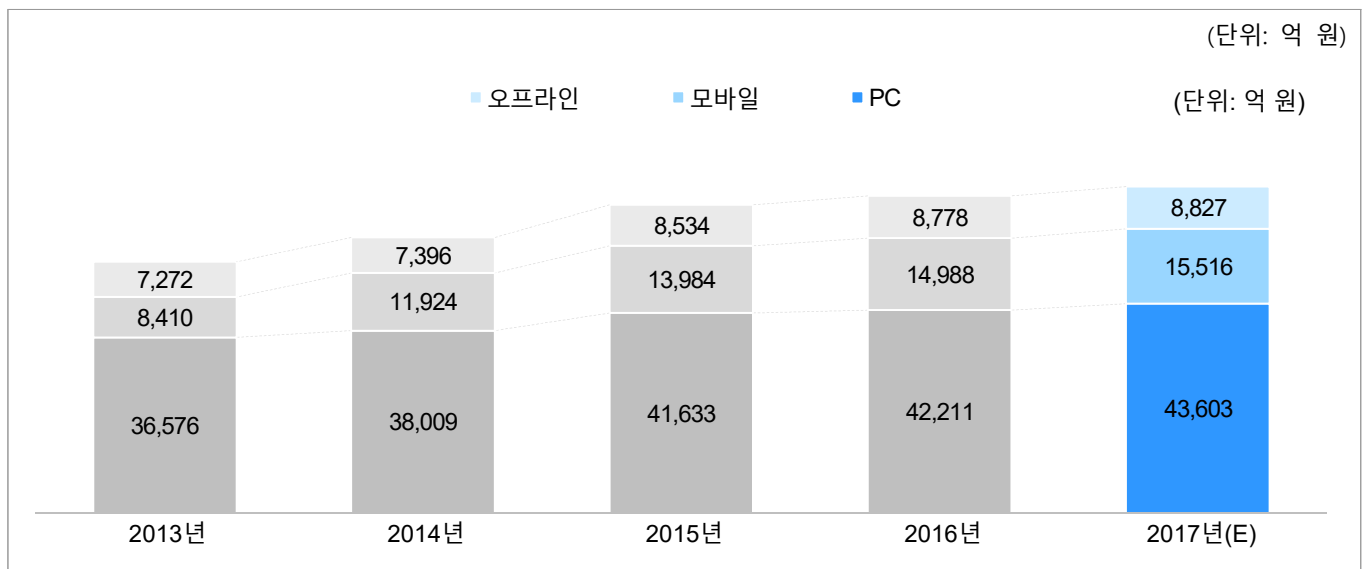
[표2-7] 데이터서비스 중분류별 시장 규모

구분	2013년		2014년		2015년		2016년		2017년(E)		증감률 '16~'17	CAGR '13~'17
	규모	비중	규모	비중	규모	비중	규모	비중	규모	비중		
데이터거래	1,708	3.3%	2,019	3.5%	2,379	3.7%	2,594	3.9%	2,709	4.0%	4.4%	12.2%
정보 제공	48,284	92.4%	51,985	90.7%	58,171	90.7%	59,854	90.7%	61,575	90.6%	2.9%	6.3%
데이터 분석 제공	2,266	4.3%	3,325	5.8%	3,601	5.6%	3,529	5.4%	3,662	5.4%	3.8%	12.7%
데이터서비스 전체	52,258	100.0%	57,329	100.0%	64,151	100.0%	65,977	100.0%	67,946	100.0%	3.0%	6.8%

(단위: 억 원)

○ 데이터서비스 기반별 시장 규모를 보면, PC 기반으로 제공하는 온라인 서비스가 2017년 4조 3,603억 원으로 전체의 64.2%, 모바일 기반 온라인 서비스는 1조 5,516억 원으로 22.8%, 오프라인은 8,827억 원으로 13.0% 순으로 나타났다.

[그림2-11] 데이터서비스 기반별 시장 규모



- 모바일 기반 데이터서비스 수요가 점차 증가함에 따라 모바일 기반의 시장은 연평균 16.5%로 높은 성장률을 나타내고 있다.

[표2-8] 데이터서비스 기반별 시장 규모

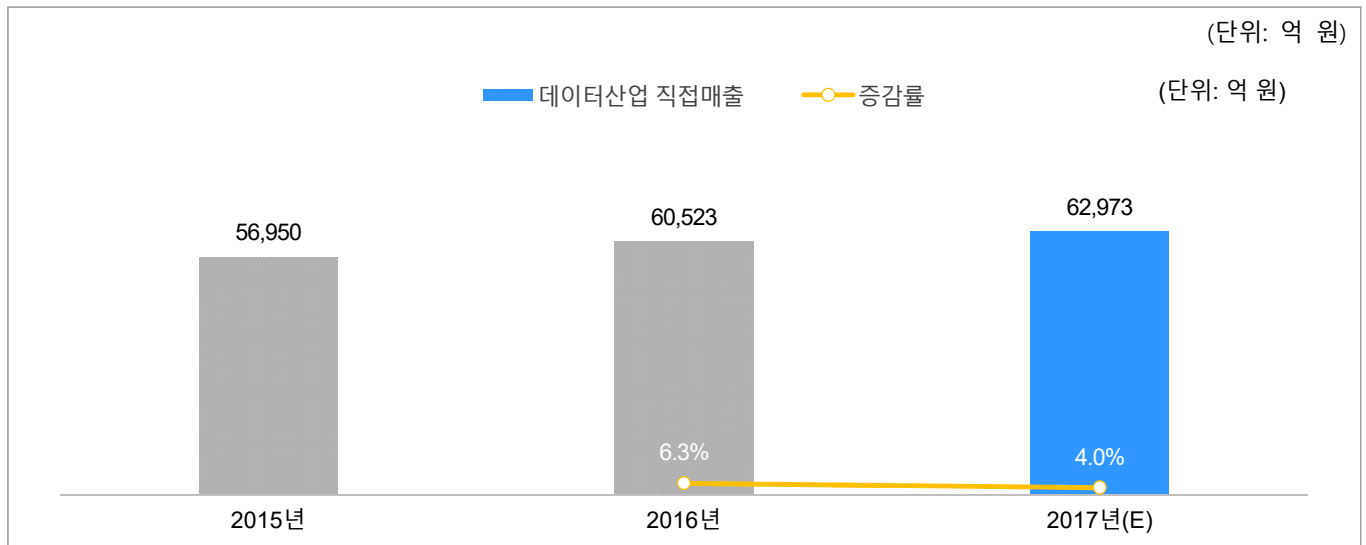
구분		2013년	2014년	2015년	2016년	2017년(E)	증감률 '16~'17	CAGR '13~'17
온라인	PC	36,576	38,009	41,633	42,211	43,603	3.3%	4.5%
		70.0%	66.3%	64.9%	64.0%	64.2%		
	모바일	8,410	11,924	13,984	14,988	15,516	3.5%	16.5%
		16.1%	20.8%	21.8%	22.7%	22.8%		
오프라인		7,272	7,396	8,534	8,778	8,827	0.6%	5.0%
		13.9%	12.9%	13.3%	13.3%	13.0%		
데이터서비스 전체		52,258	57,329	64,151	65,977	67,946	3.0%	6.8%
		100.0%	100.0%	100.0%	100.0%	100.0%		

2.2 데이터산업 직접매출 시장 규모

1) 전체 시장

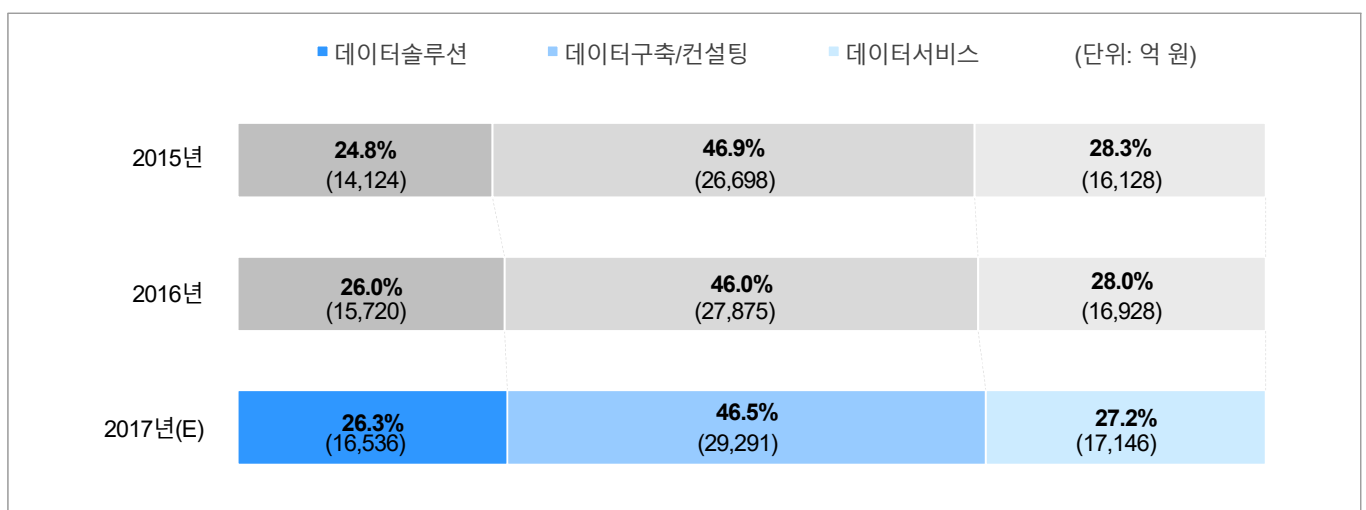
- 국내 데이터산업 직접매출 시장 규모는 2016년 6조 523억 원이며, 2017년에는 6조 2,973억 원으로 전년 대비 4.0% 성장할 것으로 나타났다.¹⁸⁾

[그림2-12] 데이터산업 직접매출 시장 규모



- 데이터산업 직접매출 시장은 기존 부문별 시장 규모에서(특히 데이터구축과 데이터서비스 부문) 데이터를 매개로 하는 광고 매출, DB시스템 구축 용역 매출 등 DB 관련 간접 매출을 제외하였다.¹⁹⁾ 그 결과 2017년 데이터구축 및 컨설팅 시장은 2조 9,291억 원, 데이터서비스 시장은 1조 7,146억 원이며, 데이터솔루션 시장은 1조 6,536억 원으로 나타났다.

[그림2-13] 데이터산업 부문별 직접매출 시장 규모 비중



18) 국내 데이터산업 직접매출 시장 규모는 데이터와 직접적으로 관련이 있는 매출(직접매출) 기준으로 적용하였음

19) 데이터산업 직접매출 시장 범위에서 데이터솔루션과 데이터컨설팅은 기존 데이터산업 규모와 동일함

- 2016년 대비 2017년 증감률은 데이터솔루션과 데이터구축/컨설팅 부문이 상대적으로 높게 나타났다.

[표2-9] 데이터산업 부문별 직접매출 시장 규모

구분	2015년		2016년		2017년(E)		증감률 '16~'17
	시장 규모	비중	시장 규모	비중	시장 규모	비중	
데이터솔루션	14,124	24.8%	15,720	26.0%	16,536	26.3%	5.2%
데이터구축/컨설팅	26,698	46.9%	27,875	46.0%	29,291	46.5%	5.1%
데이터서비스	16,128	28.3%	16,928	28.0%	17,146	27.2%	1.3%
전체	56,950	100.0%	60,523	100.0%	62,973	100.0%	4.0%

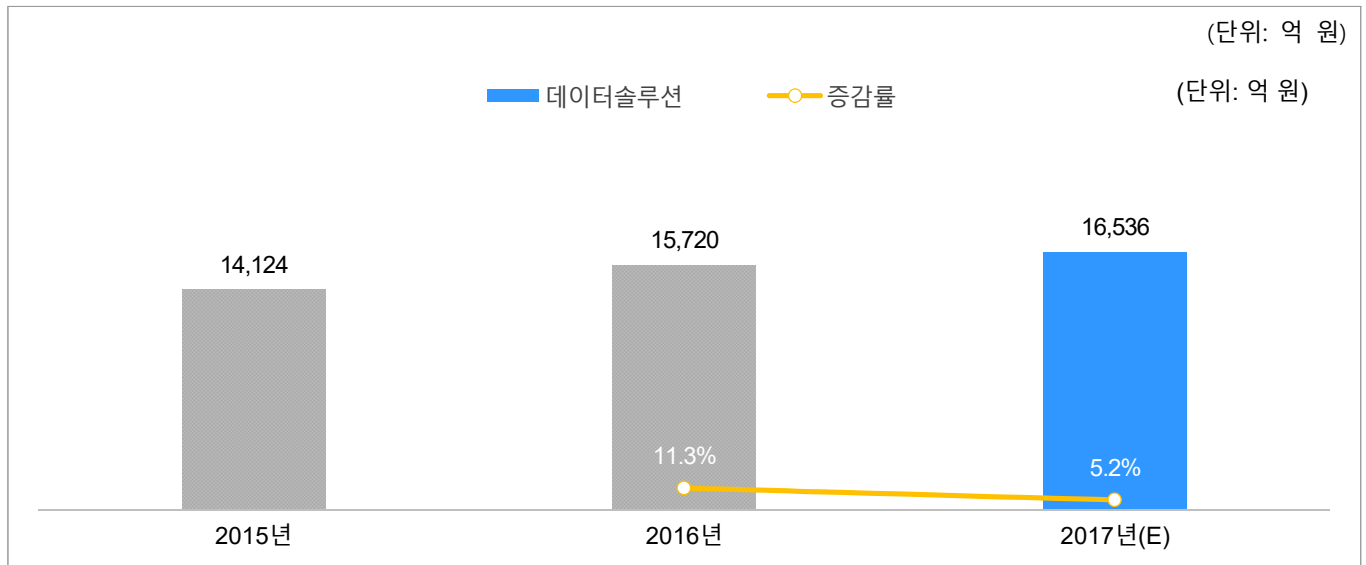
(단위: 억 원)

2) 부문별 시장

가. 데이터솔루션 시장

- 데이터솔루션 직접매출 시장 규모는 2016년 1조 5,720억 원이며, 2017년에는 1조 6,536억 원으로 전년 대비 5.2% 성장할 것으로 나타났다.²⁰⁾

[그림2-14] 데이터솔루션 직접매출 시장 규모



[표2-10] 데이터솔루션 중분류별 직접매출 시장 규모

구분	2015년		2016년		2017년(E)		증감률 '16~'17
	시장 규모	비중	시장 규모	비중	시장 규모	비중	
데이터 수집	1,115	7.9%	1,345	8.6%	1,382	8.4%	2.8%
데이터 설계	207	1.5%	225	1.4%	269	1.6%	19.6%
DBMS	5,727	40.5%	6,148	39.1%	6,192	37.5%	0.7%
데이터 관리	3,574	25.3%	4,022	25.6%	4,417	26.7%	9.8%
데이터 품질	918	6.5%	1,120	7.1%	1,127	6.8%	0.6%
데이터 분석	1,157	8.2%	1,249	7.9%	1,328	8.0%	6.3%
데이터 플랫폼	1,426	10.1%	1,611	10.3%	1,821	11.0%	13.0%
데이터솔루션 전체	14,124	100.0%	15,720	100.0%	16,536	100.0%	5.2%

(단위: 억 원)

20) 데이터산업 직접매출 시장 규모 중 데이터솔루션 부문 시장 규모는 기존 데이터솔루션 시장 규모를 그대로 유지하였음

○ 라이선스, 개발, 유지보수 영역별 규모는 다음과 같다.

[표2-11] 데이터솔루션 영역별 직접매출 시장 규모

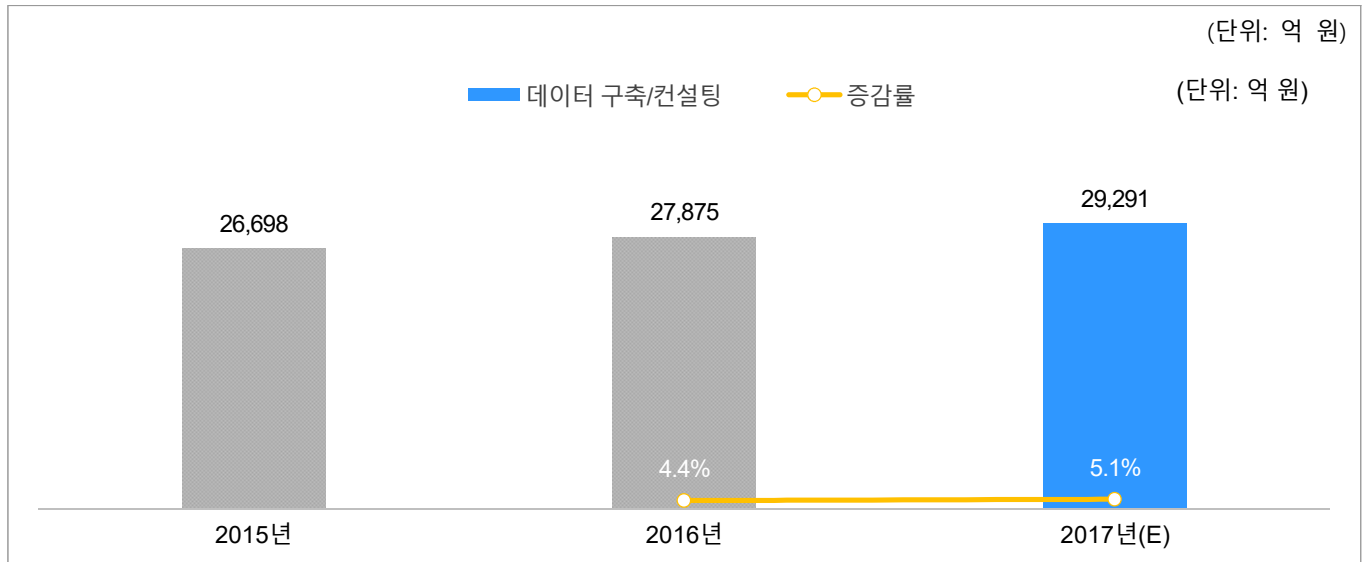
구 분		2015년	2016년	2017년(E)	증감률 '16~'17
데이터 수집	라이선스	225	277	289	4.3%
	개발	523	681	673	-1.2%
	유지보수	367	387	420	8.5%
데이터 설계	라이선스	33	42	50	19.0%
	개발	133	125	142	13.6%
	유지보수	41	58	77	32.8%
DBMS	라이선스	2,712	2,880	2,948	2.4%
	개발	2,101	2,184	2,127	-2.6%
	유지보수	914	1,084	1,117	3.0%
데이터 관리	라이선스	363	452	474	4.9%
	개발	1,796	1,902	1,987	4.5%
	유지보수	1,415	1,668	1,956	17.3%
데이터 품질	라이선스	131	179	210	17.3%
	개발	354	438	433	-1.1%
	유지보수	433	503	484	-3.8%
데이터 분석	라이선스	116	190	207	8.9%
	개발	846	830	836	0.7%
	유지보수	195	229	285	24.5%
데이터 플랫폼	라이선스	361	352	432	22.7%
	개발	823	958	1,068	11.5%
	유지보수	242	301	321	6.6%
데이터솔루션 전체	라이선스	3,941	4,372	4,610	5.4%
	개발	6,576	7,118	7,266	2.1%
	유지보수	3,607	4,230	4,660	10.2%

(단위: 억 원)

나. 데이터구축/컨설팅 시장

- 데이터구축/컨설팅 직접매출 시장 규모는 2016년 2조 7,875억 원이며, 2017년은 2조 9,291억 원으로 전년 대비 5.1% 상승할 것으로 조사되었다.²¹⁾

[그림2-15] 데이터구축/컨설팅 직접매출 시장 규모



[표2-12] 데이터 구축/컨설팅 직접매출 시장 규모

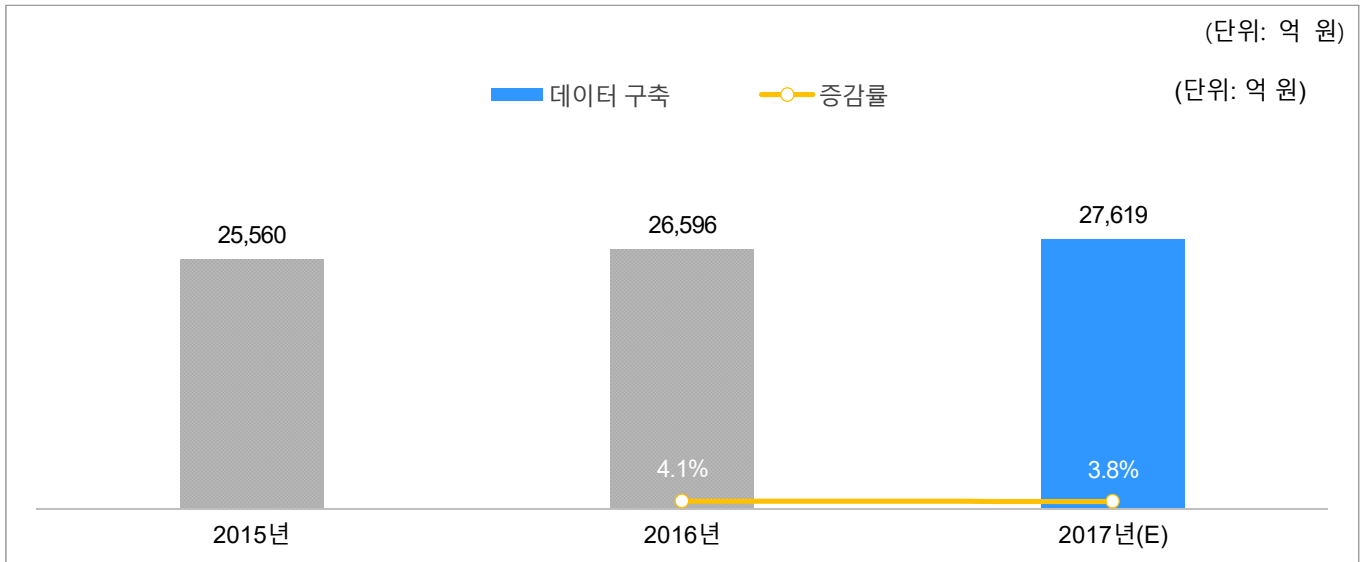
구분	2015년		2016년		2017년(E)		증감률 '16~'17
	시장 규모	비중	시장 규모	비중	시장 규모	비중	
데이터 구축	25,560	95.7%	26,596	95.4%	27,619	94.3%	3.8%
데이터 컨설팅	1,138	4.3%	1,279	4.6%	1,672	5.7%	30.7%
데이터 구축/컨설팅 전체	26,698	100.0%	27,875	100.0%	29,291	100.0%	5.1%

(단위: 억 원)

21) 데이터산업 직접매출 시장 규모 중 데이터구축 부문 시장 규모는 기존 데이터구축 시장 규모에서 DB시스템 구축 등 DB기반의 간접매출을 제외하고 데이터와 직접적으로 관련 있는 매출을 중심으로 산출한 결과이며, 컨설팅 부문은 그대로 유지하였음

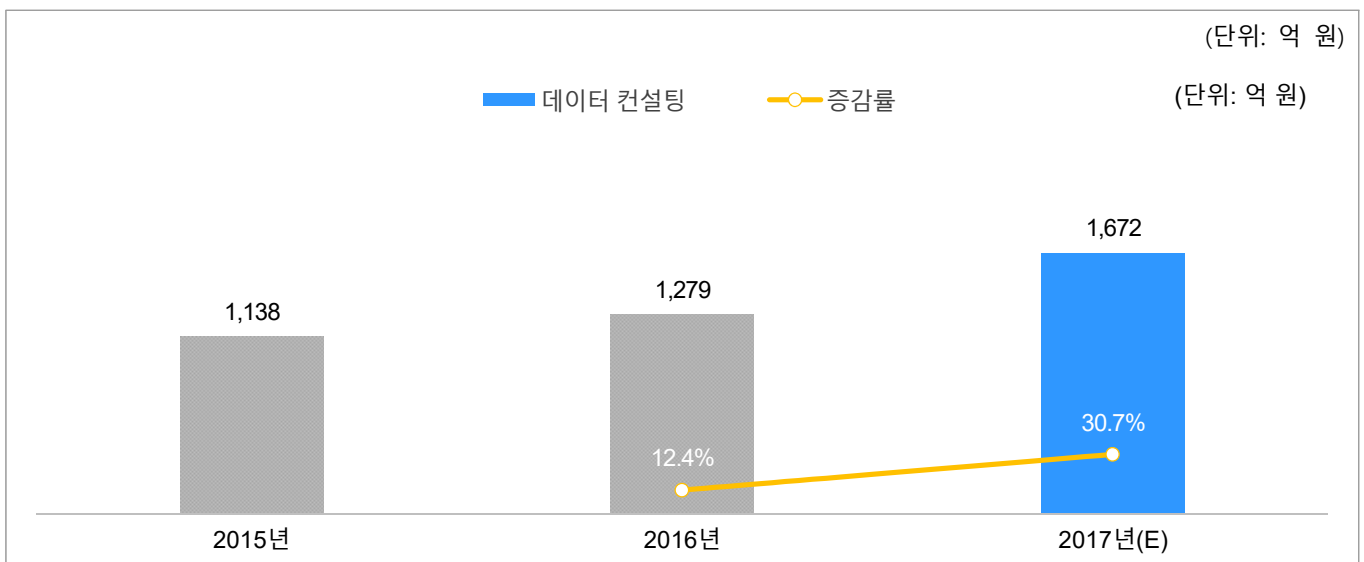
- 데이터구축 직접매출 시장 규모는 2016년 2조 6,596억 원이며, 2017년에는 전년 대비 3.8% 증가한 2조 7,619억 원으로 전년과 비슷한 수준을 유지할 것으로 나타났다.

[그림2-16] 데이터구축 직접매출 시장 규모



- 데이터컨설팅 직접매출 시장은 2016년 1,279억 원이며, 2017년에는 1,672억 원으로 2016년 대비 30.7% 성장할 것으로 조사되었다.

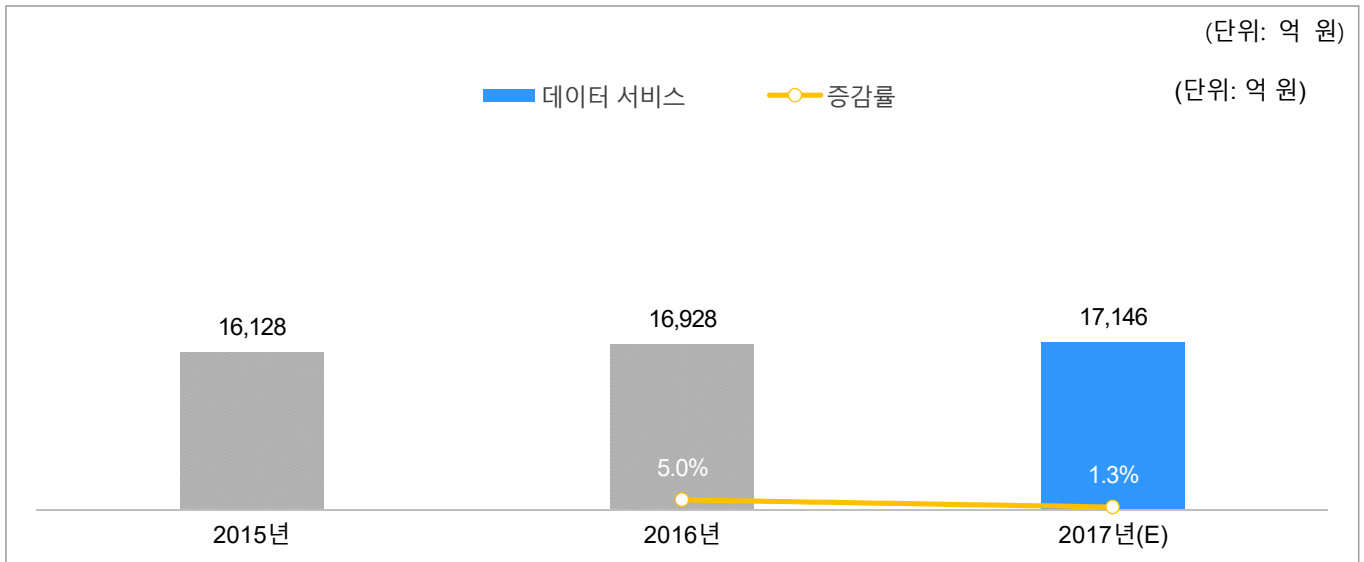
[그림2-17] 데이터컨설팅 직접매출 시장 규모



다. 데이터서비스 시장

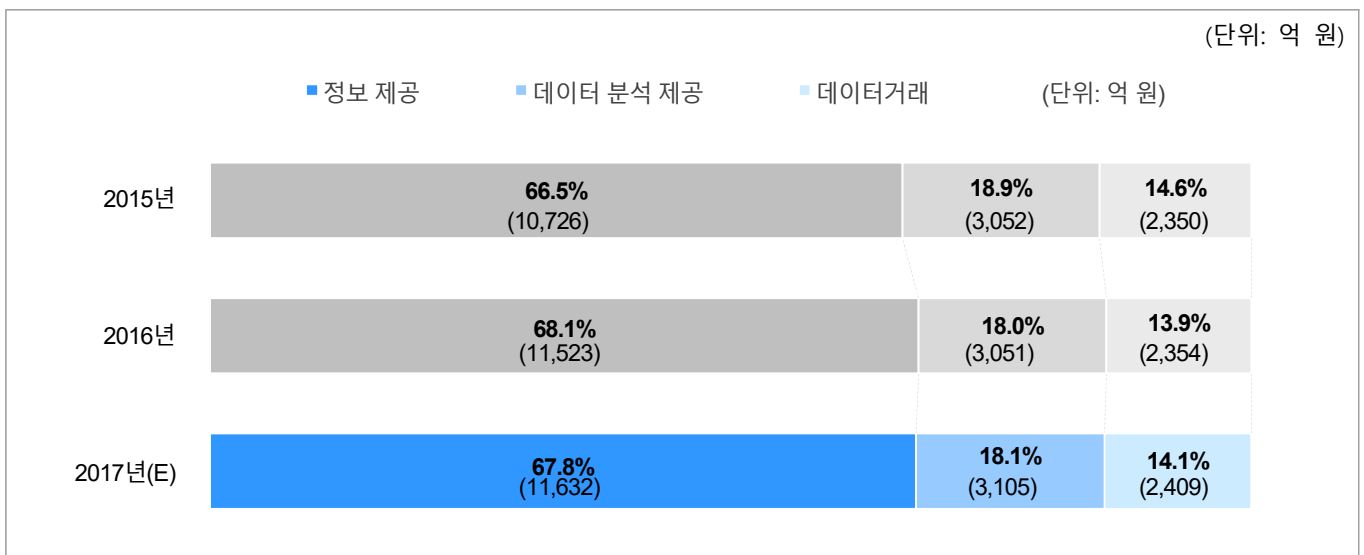
- 데이터서비스 직접매출 시장은 2016년 1조 6,928억 원이며, 2017년에는 1조 7,146억 원으로 전년 대비 1.3% 증가할 것으로 조사되었다.²²⁾

[그림2-18] 데이터서비스 직접매출 시장 규모



- 데이터서비스 중분류별 직접매출 시장 규모를 보면, 정보 제공 부문이 절반 이상을 차지하며, 2016년은 1조 1,523억 원, 2017년은 전년 대비 0.9% 증가한 1조 1,632억 원 수준으로 나타났다.
- 2017년 데이터 거래 시장은 2016년 대비 2.3% 증가한 2,409억 원이며, 데이터 분석 제공 시장은 2016년 대비 1.8% 증가한 3,105억 원으로 나타났다.

[그림2-19] 데이터서비스 중분류별 직접매출 시장 규모 비중



22) 데이터산업 직접매출 시장 규모 중 데이터서비스 부문 시장 규모는 기존 데이터서비스 시장 규모에서 광고매출, DB기반의 정보서비스 등의 데이터 간접매출을 제외하고 데이터와 직접적으로 관련이 있는 매출을 중심으로 산출한 결과임

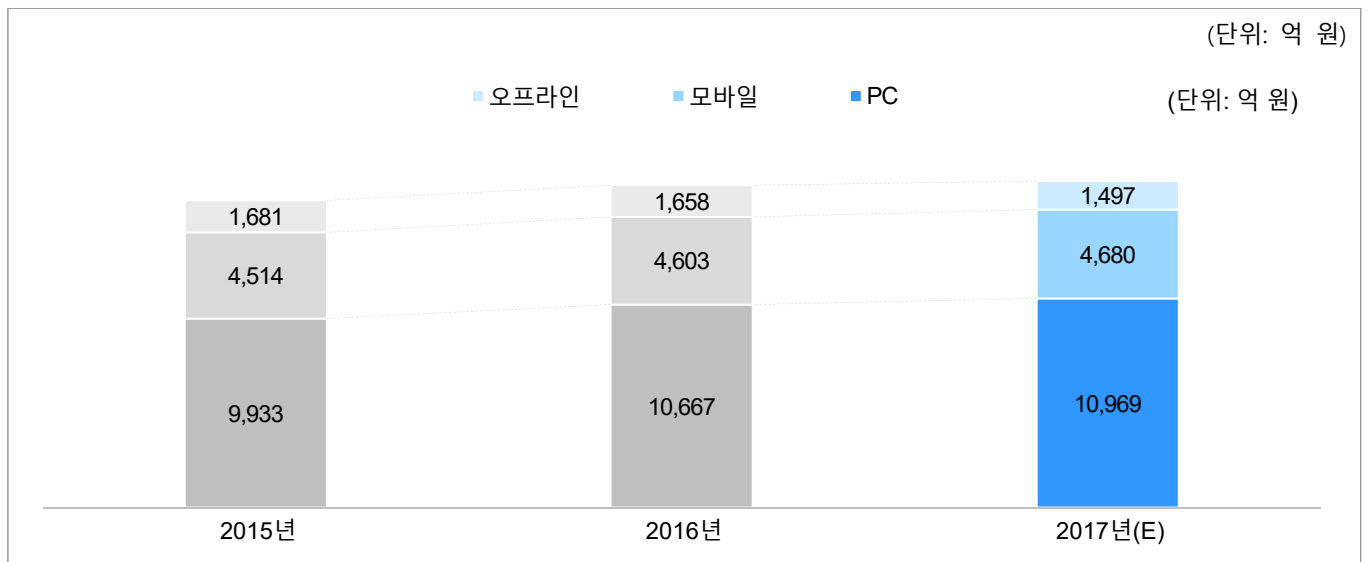
[표2-13] 데이터서비스 중분류별 직접매출 시장 규모

구분	2015년		2016년		2017년(E)		증감률 '16~'17
	시장 규모	비중	시장 규모	비중	시장 규모	비중	
데이터거래	2,350	14.6%	2,354	13.9%	2,409	14.1%	2.3%
정보 제공	10,726	66.5%	11,523	68.1%	11,632	67.8%	0.9%
데이터 분석 제공	3,052	18.9%	3,051	18.0%	3,105	18.1%	1.8%
데이터서비스 전체	16,128	100.0%	16,928	100.0%	17,146	100.0%	1.3%

(단위: 억 원)

- 서비스 기반별로 살펴보면, PC 기반으로 제공하는 온라인 서비스가 2017년 1조 969억 원으로 전체의 64.0%를 차지할 것으로 보이며, 모바일 기반 온라인 서비스는 4,680억 원으로 27.3%, 오프라인은 1,497억 원으로 8.7% 순으로 나타났다.

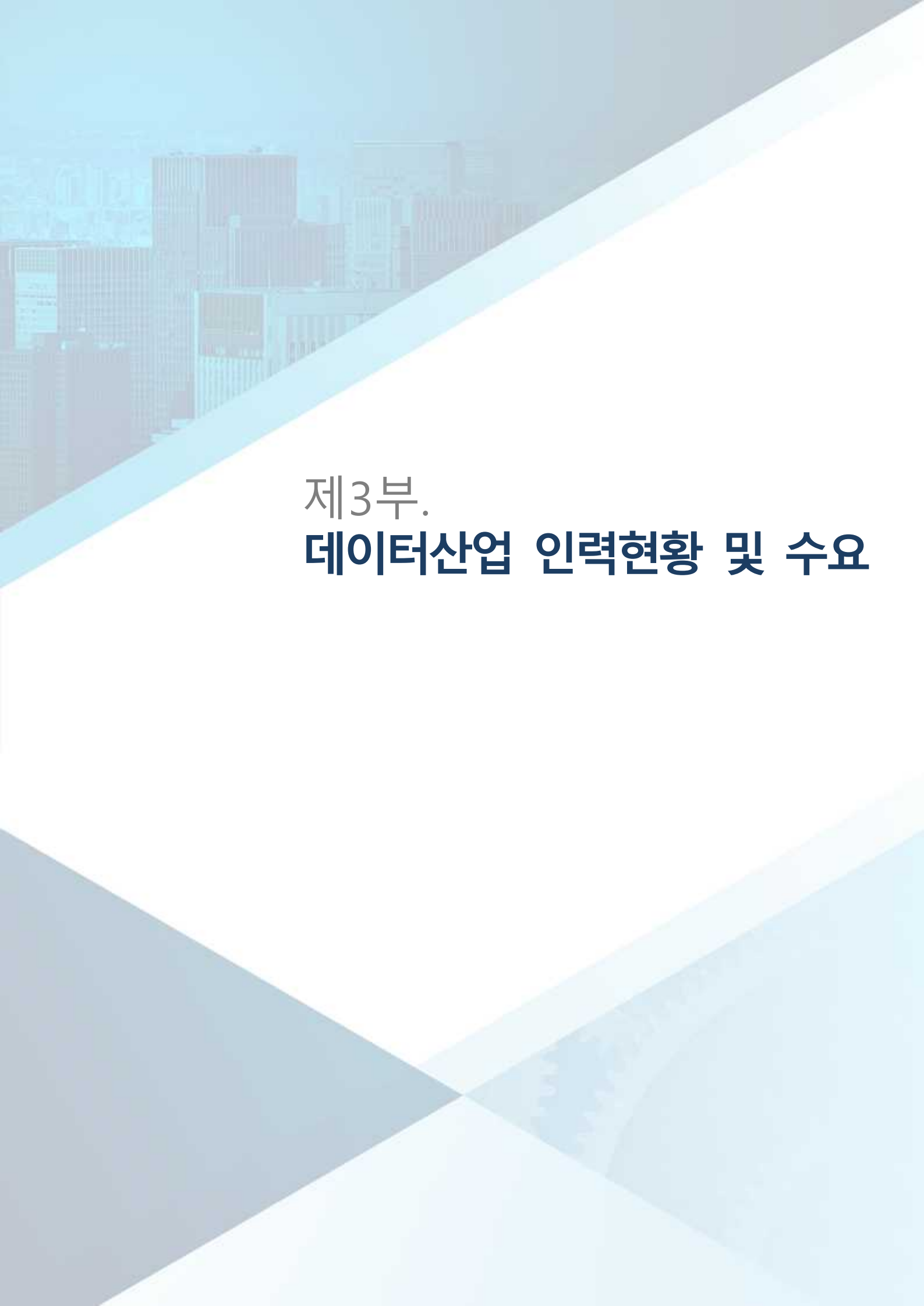
[그림2-20] 데이터서비스 기반별 직접매출 시장 규모



[표2-14] 데이터서비스 기반별 직접매출 시장 규모

구분		2015년		2016년		2017년(E)		증감률 '16~'17
		시장 규모	비중	시장 규모	비중	시장 규모	비중	
온라인	PC	9,933	61.6%	10,667	63.0%	10,969	64.0%	2.8%
	모바일	4,514	28.0%	4,603	27.2%	4,680	27.3%	1.7%
오프라인		1,681	10.4%	1,658	9.8%	1,497	8.7%	-9.7%
데이터서비스 전체		16,128	100.0%	16,928	100.0%	17,146	100.0%	1.3%

(단위: 억 원)

The background features a stylized city skyline with various skyscrapers in shades of blue and grey. A large, light blue diagonal band runs from the top left towards the bottom right. In the bottom right corner, there is a faint, light blue gear-like pattern.

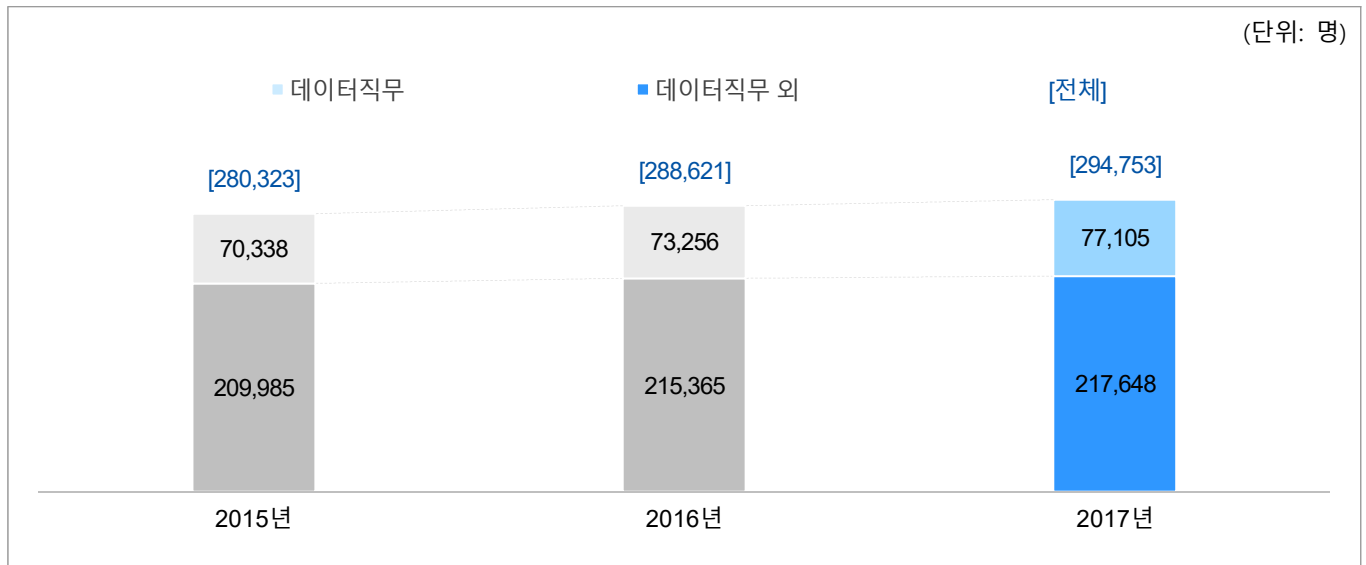
제3부. 데이터산업 인력현황 및 수요

3.1 인력 현황

1) 데이터산업 인력 현황

- 2017년 데이터산업에 종사하고 있는 인력은 총 29만 4,753명으로 전년 대비 2.1% 증가하였으며, 이 중 데이터직무 인력은 7만 7,105명으로 전년 대비 5.3% 증가한 것으로 나타났다.

[그림3-1] 데이터산업 인력 현황



- 데이터직무 인력이 데이터직무 외 인력보다 전년 대비 증가율이 높게 나타났으며, 이는 데이터 분석가, 데이터 개발자 인력 증가로 인한 결과로 분석된다.

[표3-1] 데이터산업 인력 현황

구분	2015년		2016년		2017년		증감률(%) '16~'17
	인력 규모 (명)	비중 (%)	인력 규모 (명)	비중 (%)	인력 규모 (명)	비중 (%)	
데이터직무	70,338	25.1%	73,256	25.4%	77,105	26.2%	5.3%
데이터직무 외	209,985	74.9%	215,365	74.6%	217,648	73.8%	1.1%
전체	280,323	100.0%	288,621	100.0%	294,753	100.0%	2.1%

(단위: 명, %)

- 데이터산업 부문별 데이터직무 인력을 보면, 데이터솔루션 분야의 인력이 11.0% 증가하였으며, 구축/컨설팅 분야의 데이터직무 인력이 전년 대비 6.0% 증가하였다.

[표3-2] 데이터산업 내 부문별 데이터직무 인력 현황

구분	2015년		2016년		2017년		증감률(%) '16~'17
	인력 규모 (명)	비중 (%)	인력 규모 (명)	비중 (%)	인력 규모 (명)	비중 (%)	
데이터솔루션	8,886	12.6%	9,272	12.7%	10,291	13.3%	11.0%
데이터구축/컨설팅	34,323	48.8%	35,404	48.4%	37,516	48.7%	6.0%
구축	29,304	41.7%	29,941	40.9%	31,693	41.1%	5.9%
컨설팅	5,019	7.1%	5,463	7.5%	5,823	7.6%	6.6%
데이터서비스	27,129	38.6%	28,580	39.0%	29,298	38.0%	2.5%
전체	70,338	100.0%	73,256	100.0%	77,105	100.0%	5.3%

(단위: 명, %)

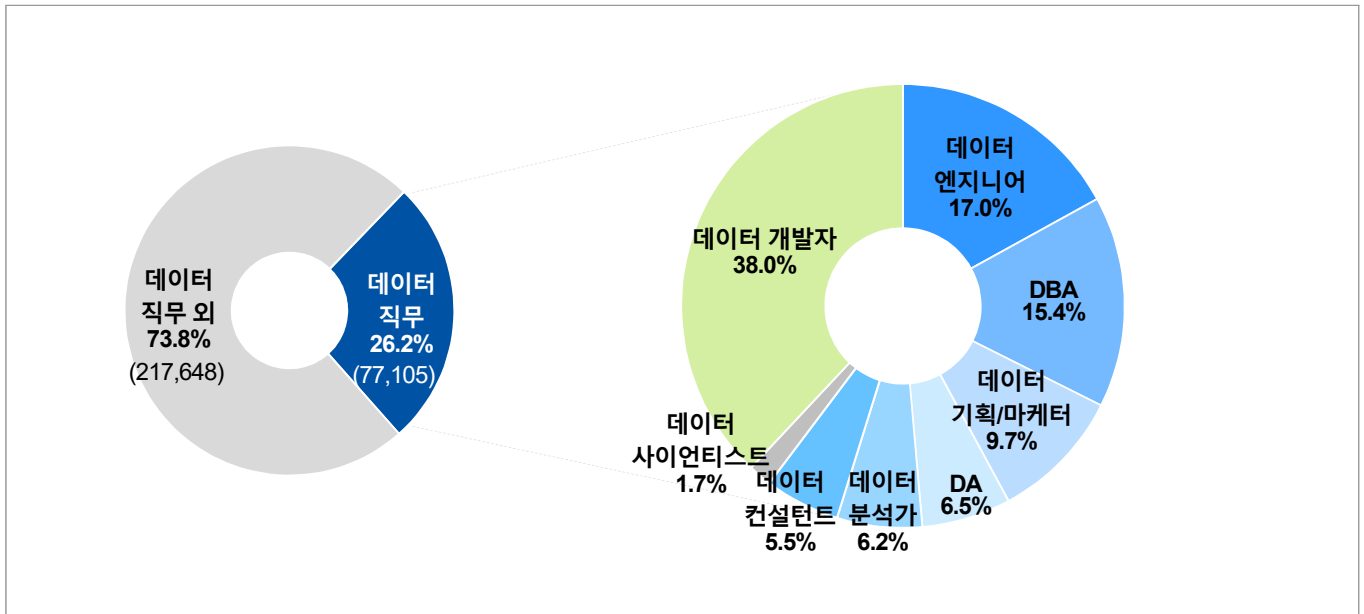
※ 이후 데이터구축/컨설팅 분야 데이터직무 인력은 편의상 데이터구축과 데이터컨설팅으로 구분해 표시하겠음

[표3-3] 데이터직무 분류표

데이터직무명	세부 내용
DA	- 데이터를 기반으로 IT 정책, 표준화, 구조, 설계 및 이행을 하는 직무 - 개념적, 논리적, 물리적 데이터 설계 수행
데이터 개발자	- 조직업무 기반 IT시스템 구축에서 DB 및 데이터를 이용하여 프로그래밍 하는 직무 - Hadoop, NoSQL, MapReduce 등의 기술을 활용해 빅데이터 처리, LOD 구축 등 데이터 관련 프로그램 개발 포함
데이터 엔지니어	DBMS, Hadoop, NoSQL 등 데이터 관련 제품에 대한 기술지원자, 제품 개발자, 유지보수 등의 직무
데이터 분석가	- 정형, 비정형 데이터 등 다양한 데이터를 식별, 관리, 조작, 분석하여 기업 경영의 의사결정에 필요한 자료를 만들어내는 직무 - 통계, 머신러닝, 텍스트마이닝 기반 데이터 분석, 분석결과 시각화 등 포함
DBA	DB관리체계와 자료를 검토, 개선하고 DB의 구성, 변경, 용량, 성능, 가용성, 보안, 장애, 문제관리 등 운영시스템의 관리 직무
데이터 사이언티스트	조직 내외부 데이터의 관리/활용/분석 체계를 만들고, 분석을 통해 프로세스 혁신 및 신제품 개발, 마케팅 전략 결정 등의 의사결정을 이끌어내는 직무
데이터 컨설턴트	- 성능튜닝, 데이터아키텍처, 문제해결 등을 총칭하는 컨설팅사의 직무 - 빅데이터 분석을 토대로 기업이 앞으로 나아갈 방향, 해결책 등을 제시하는 업무 포함
데이터 기획·마케터	(빅)데이터 관련 제품/서비스 기획, 기술영업 및 마케팅 직무

- 2017년 데이터산업에 종사하는 데이터직무 인력 총 7만 7,105명 중 데이터 개발자가 전체의 38.0%로 가장 큰 비중을 차지하며, 이어서 데이터 엔지니어가 17.0%, DBA가 15.4%, 데이터 기획·마케터 9.7% 순으로 나타났다.

[그림3-2] 데이터산업 내 데이터직무별 인력 비중



[표3-4] 2017년 데이터산업 내 데이터직무별 인력 현황

구분	데이터솔루션	데이터구축	데이터컨설팅	데이터서비스	전체
DA	618 6.0%	3,392 10.7%	526 9.0%	489 1.7%	5,025 6.5%
데이터 개발자	4,429 43.0%	13,702 43.2%	2,188 37.6%	8,948 30.5%	29,267 38.0%
데이터 엔지니어	1,468 14.3%	6,529 20.6%	696 11.9%	4,434 15.1%	13,127 17.0%
데이터 분석가	629 6.1%	1,638 5.2%	335 5.8%	2,190 7.5%	4,792 6.2%
DBA	702 6.8%	3,732 11.8%	388 6.7%	7,029 24.0%	11,851 15.4%
데이터 사이언티스트	218 2.1%	410 1.3%	173 3.0%	482 1.7%	1,283 1.7%
데이터 컨설턴트	1,530 14.9%	1,376 4.3%	1,231 21.1%	126 0.4%	4,263 5.5%
데이터 기획·마케터	697 6.8%	914 2.9%	286 4.9%	5,600 19.1%	7,497 9.7%
전체	10,291 100.0%	31,693 100.0%	5,823 100.0%	29,298 100.0%	77,105 100.0%

(단위: 명, %)

- 빅데이터 관련 인력은 데이터직무 인력 중에서 빅데이터 관련 기술을 보유한 인력을 말한다.²³⁾
데이터산업의 데이터직무 인력 중 빅데이터 인력은 총 6,324명이며, 전체의 8.2%를 차지하였다.

[표3-5] 데이터산업 내 데이터직무 중 빅데이터 관련 인력 현황

구분	데이터산업 내 데이터직무 인력				데이터산업 내 데이터직무 중 빅데이터 인력 비중(%)
	전체		빅데이터 관련 인력		
	규모(명)	비중(%)	규모(명)	비중(%)	
DA	5,025	6.5%	-	-	-
데이터 개발자	29,267	38.0%	1,866	29.5%	6.4%
데이터 엔지니어	13,127	17.0%	1,081	17.1%	8.2%
데이터 분석가	4,792	6.2%	381	6.0%	8.0%
DBA	11,851	15.4%	-	-	-
데이터 사이언티스트	1,283	1.7%	1,283	20.3%	100.0%
데이터 컨설턴트	4,263	5.5%	1,369	21.7%	32.1%
데이터 기획·마케터	7,497	9.7%	344	5.4%	4.6%
전체	77,105	100.0%	6,324	100.0%	8.2%

(단위: 명, %)

※ 빅데이터 관련 인력 : 데이터직무 인력 중 빅데이터 기술을 보유한 인력으로 데이터직무 인력 수에 포함됨

※ DA와 DBA는 조사 제외 대상이며, 데이터 사이언티스트 직무는 100% 빅데이터 관련 인력으로 간주

- 현재 데이터직무 중 빅데이터 인력 비중을 살펴보면, 데이터솔루션 분야에서 빅데이터 인력 비중이 가장 높게 나타났다. 특히 데이터 엔지니어 비중이 높았으며, 데이터 기획·마케터 비중도 비교적 높게 나타났다.
- 데이터컨설팅 분야에서는 빅데이터 분석가와 데이터 컨설턴트의 비중이 비교적 높게 나타났다.

23) 빅데이터 관련 기술 예시는 다음과 같음

- ① Hadoop, NoSQL, MapReduce, Spark 등 빅데이터를 처리하거나 프로그램을 개발하기 위한 관련 기술 보유
- ② 두 개 이상의 데이터 포맷(예: 정형, 비정형)과 데이터 소스(예: 내부, 외부)를 활용하는 빅데이터 분석 기술 보유
- ③ 실시간으로 스트리밍 되는 데이터 분석 기술 보유
- ④ 빅데이터 관련 제품 및 서비스 기획 능력 보유
- ⑤ 빅데이터 분석을 통한 문제해결, 의사결정 도출 및 제시 능력
- ⑥ 기타 빅데이터 분석, 처리, 개발 등을 위한 관련 기술 보유 등

[표3-6] 데이터산업 내 데이터직무 중 빅데이터 관련 인력 비중

구분	데이터산업 내 데이터직무 중 빅데이터 관련 인력 비중				
	데이터솔루션	데이터구축	데이터컨설팅	데이터서비스	전체
데이터 개발자	20.2%	3.0%	16.0%	2.4%	6.4%
데이터 엔지니어	56.5%	2.1%	8.2%	1.3%	8.2%
데이터 분석가	5.9%	8.2%	58.5%	0.6%	8.0%
데이터 사이언티스트	100.0%	100.0%	100.0%	100.0%	100.0%
데이터 컨설턴트	17.7%	35.2%	45.7%	41.3%	32.1%
데이터 기획·마케터	37.9%	2.8%	5.6%	0.7%	4.6%
전체	24.4%	5.0%	23.3%	2.9%	8.2%

(단위: %)

- 데이터솔루션, 데이터구축, 데이터컨설팅, 데이터 서비스 분야 순으로 빅데이터 인력이 많으며, 데이터컨설팅 분야는 빅데이터 관련 데이터 컨설턴트 인력 비중이 41.5%로 특히 높게 나타났다.

[표3-7] 데이터산업 내 부문별 빅데이터 관련 인력 현황

구분	데이터산업 내 부문별 빅데이터 관련 인력 현황				
	데이터솔루션	데이터구축	데이터컨설팅	데이터서비스	전체
데이터 개발자	896	409	350	211	1,866
	35.6%	25.6%	25.8%	24.7%	29.5%
데이터 엔지니어	829	137	57	58	1,081
	32.9%	8.6%	4.2%	6.8%	17.1%
데이터 분석가	37	134	196	14	381
	1.5%	8.4%	14.5%	1.6%	6.0%
데이터 사이언티스트	218	410	173	482	1,283
	8.7%	25.6%	12.8%	56.4%	20.3%
데이터 컨설턴트	271	484	562	52	1,369
	10.8%	30.2%	41.5%	6.1%	21.6%
데이터 기획·마케터	264	26	16	38	344
	10.5%	1.6%	1.2%	4.4%	5.5%
전체	2,515	1,600	1,354	855	6,324
	100.0%	100.0%	100.0%	100.0%	100.0%

(단위: 명, %)

2) 전 산업의 데이터직무 인력 현황

- 일반산업을 포함한 전 산업의 2017년 데이터직무 인력은 총 10만 9,320명으로 조사되었다.²⁴⁾ 데이터산업은 전년 대비 5.3% 증가한 7만 7,105명이며, 일반산업은 10.6% 증가한 3만 2,215명으로 나타났다.

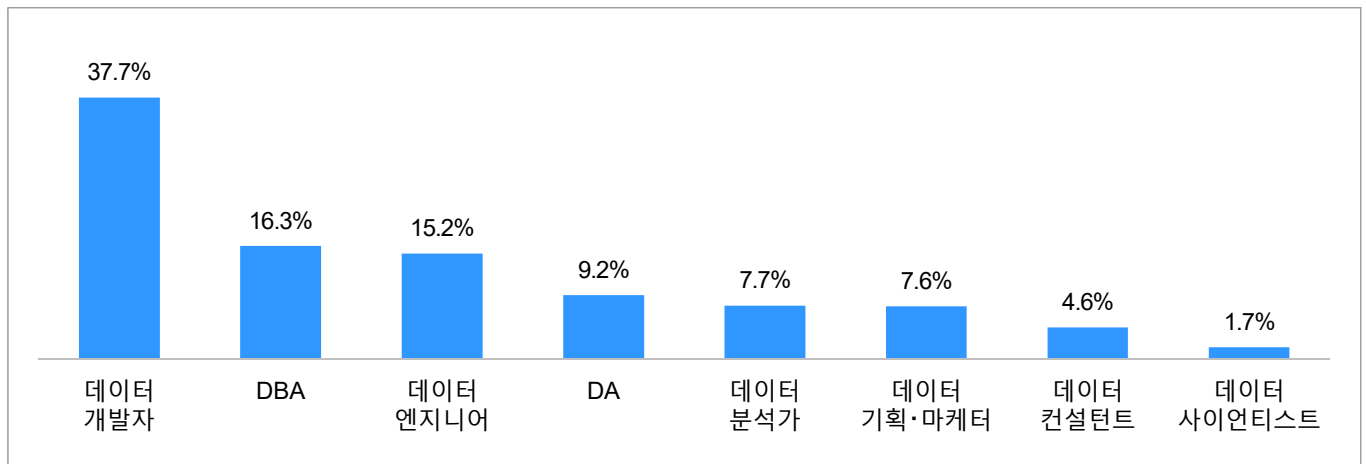
[표3-8] 전 산업 내 데이터직무 인력 현황

구분	2015년		2016년		2017년		증감률(%) '16~'17
	인력 규모 (명)	비중 (%)	인력 규모 (명)	비중 (%)	인력 규모 (명)	비중 (%)	
데이터산업	70,338	70.0%	73,256	71.6%	77,105	70.5%	5.3%
일반산업	30,102	30.0%	29,119	28.4%	32,215	29.5%	10.6%
전 산업	100,440	100.0%	102,375	100.0%	109,320	100.0%	6.8%

(단위: 명, %)

- 데이터직무별로는 데이터산업과 일반산업 모두 데이터 개발자 비중이 가장 높은 것으로 나타났다.

[그림3-3] 전 산업 내 데이터직무 비중



24) 일반산업 모집단 및 조사표본 수는 다음과 같음

구분	모집단 기업 수 (종사자 100명 이상)	2016년		2017년	
		조사표본 수	비중	조사표본 수	비중
공공	1,157	202	23.1%	164	18.7%
금융	399	43	4.9%	44	5.0%
유통/서비스	1,916	146	16.7%	186	21.3%
제조	4,195	357	40.9%	353	40.3%
의료	417	64	7.3%	74	8.5%
통신/미디어	1,325	62	7.1%	54	6.2%
일반산업 전체	9,409	874	100.0%	875	100.0%

※ 2017년 일반산업 결과는 '2017년 빅데이터 시장 현황조사' 데이터를 활용한 결과임

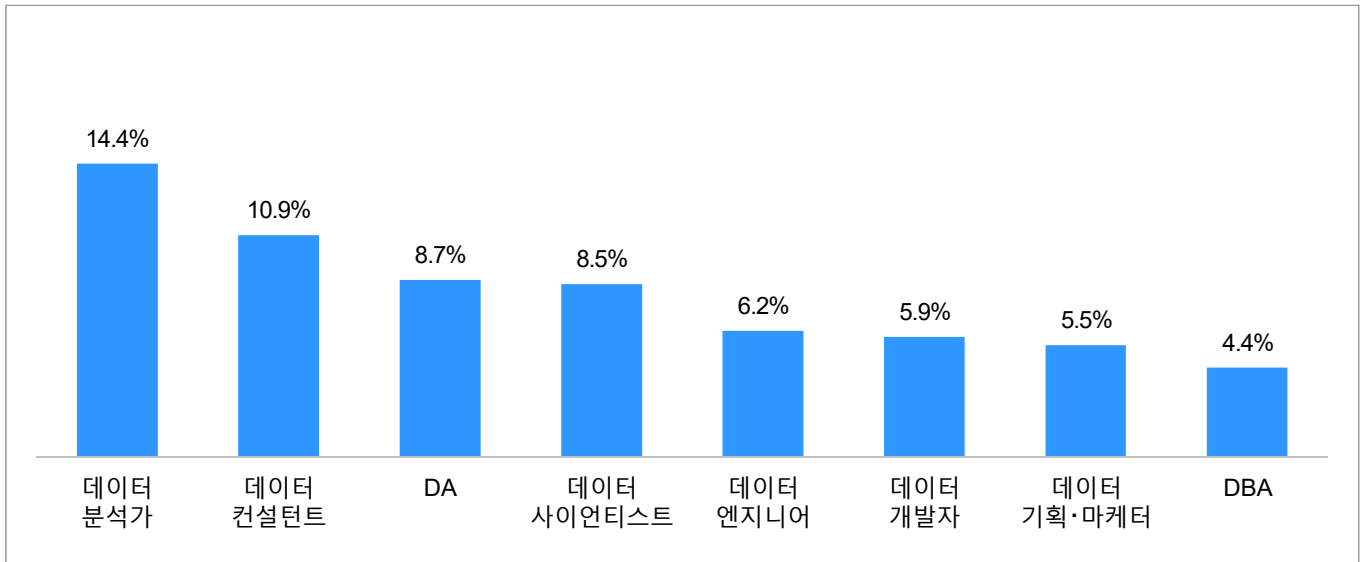
[표3-9] 2017년 전 산업 내 데이터직무별 인력 현황

구분	데이터산업	일반산업	전 산업
DA	5,025	5,046	10,071
	6.5%	15.7%	9.2%
데이터 개발자	29,267	11,987	41,254
	38.0%	37.2%	37.7%
데이터 엔지니어	13,127	3,507	16,634
	17.0%	10.9%	15.2%
데이터 분석가	4,792	3,606	8,398
	6.2%	11.2%	7.7%
DBA	11,851	6,012	17,863
	15.4%	18.6%	16.3%
데이터 사이언티스트	1,283	520	1,803
	1.7%	1.6%	1.7%
데이터 컨설턴트	4,263	741	5,004
	5.5%	2.3%	4.6%
데이터 기획·마케터	7,497	796	8,293
	9.7%	2.5%	7.6%
전체	77,105	32,215	109,320
	100.0%	100.0%	100.0%

(단위: 명, %)

- 2016년과 비교하면 데이터 분석가의 증가율은 14.4%이며, 다음으로 데이터 컨설턴트, DA, 데이터 사이언티스트 등의 순으로 나타났다.

[그림3-4] 2017년 전년대비 전 산업 내 데이터직무별 증감률



[표3-10] 2015년-2017년 전 산업 내 데이터직무별 인력 현황 비교

구분	2015년		2016년		2017년		증감률(%) '16~'17
	인력 규모 (명)	비중 (%)	인력 규모 (명)	비중 (%)	인력 규모 (명)	비중 (%)	
DA	7,065	7.0%	9,267	9.1%	10,071	9.2%	8.7%
데이터 개발자	41,205	41.0%	38,948	38.0%	41,254	37.7%	5.9%
데이터 엔지니어	14,013	14.0%	15,670	15.3%	16,634	15.2%	6.2%
데이터 분석가	5,120	5.1%	7,339	7.2%	8,398	7.7%	14.4%
DBA	17,427	17.4%	17,116	16.7%	17,863	16.3%	4.4%
데이터 사이언티스트	1,343	1.3%	1,662	1.6%	1,803	1.7%	8.5%
데이터 컨설턴트	4,351	4.3%	4,513	4.4%	5,004	4.6%	10.9%
데이터 기획·마케터	9,916	9.9%	7,860	7.7%	8,293	7.6%	5.5%
전체	100,440	100.0%	102,375	100.0%	109,320	100.0%	6.8%

(단위: 명, %)

※ 2015년과 2016년 데이터직무 분류를 조정하여 재구성한 표임

○ 데이터기술 등급 기준은 다음과 같다.

[표3-11] 데이터기술 등급 기준

기술등급	자격기준
초급	• 정보처리기사, 정보처리산업기사 등 동급 Data 관련 자격 취득자
	• 전문학사 및 학사학위 보유자
	• 고등학교 졸업 후 3년 이상 경력
중급	• 기사자격 또는 Data 관련 자격 취득 후 경력 3년 이상 (Data 관련 자격 : 데이터아키텍처전문가, SQL개발자, OCP, 데이터분석전문가, MCDBA, CSP, MySQL, DB Administrator, DB2 for Linux and Windows)
	• 산업기사 자격 취득 후 7년 이상 경력 • 석사의 경우 기사자격 또는 Data 관련 자격 취득 후 경력 2년 이상
고급	• 정보처리기사자격 또는 Data관련 자격 취득 후 3년 이상 경력
	• 정보관리 기술사, 컴퓨터시스템응용 기술사
	• 중급 이후 3년 이상 경력
	• Data 고급 자격 취득자 (Data고급자격 : 데이터아키텍처전문가, SQL전문가, OCM, 데이터 분석 전문가)
	• 박사의 경우 기사 및 Data관련 자격 취득자

○ 기술등급별 인력 현황을 살펴보면, 대체로 중급, 고급인력이 많았으며, 데이터 컨설턴트, 데이터 분석가, DBA 직무에서는 기술등급 중 고급인력이 가장 많은 것으로 나타났다.

[표3-12] 전 산업 내 데이터직무 기술등급별 인력 현황

구분		데이터 솔루션	데이터 구축	데이터 컨설팅	데이터 서비스	일반산업	전 산업
DA	초급	61	541	79	70	870	1,621
	중급	242	1,731	116	177	2,408	4,674
	고급	315	1,120	331	242	1,768	3,776
	계	618	3,392	526	489	5,046	10,071
데이터 개발자	초급	1,377	4,101	513	1,251	2,377	9,619
	중급	1,637	5,152	907	3,877	5,377	16,950
	고급	1,415	4,449	768	3,820	4,233	14,685
	계	4,429	13,702	2,188	8,948	11,987	41,254
데이터 엔지니어	초급	314	1,736	128	620	763	3,561
	중급	645	2,479	343	2,100	1,773	7,340
	고급	509	2,314	225	1,714	971	5,733
	계	1,468	6,529	696	4,434	3,507	16,634
데이터 분석가	초급	86	266	62	243	571	1,228
	중급	240	560	120	1,062	1,578	3,560
	고급	303	812	153	885	1,457	3,610
	계	629	1,638	335	2,190	3,606	8,398
DBA	초급	117	685	38	486	1,247	2,573
	중급	309	1,327	172	2,793	2,779	7,380
	고급	276	1,720	178	3,750	1,986	7,910
	계	702	3,732	388	7,029	6,012	17,863
데이터 사이언티스트	초급	19	41	63	34	71	228
	중급	138	92	2	333	274	839
	고급	61	277	108	115	175	736
	계	218	410	173	482	520	1,803
데이터 컨설턴트	초급	165	136	160	20	75	556
	중급	575	554	524	58	302	2,013
	고급	790	686	547	48	364	2,435
	계	1,530	1,376	1,231	126	741	5,004
데이터 기획·마케터	초급	108	148	41	884	167	1,348
	중급	301	440	160	2,759	353	4,013
	고급	288	326	85	1,957	276	2,932
	계	697	914	286	5,600	796	8,293
전체	초급	2,247	7,654	1,084	3,608	6,141	20,734
	중급	4,087	12,335	2,344	13,159	14,844	46,769
	고급	3,957	11,704	2,395	12,531	11,230	41,817
	계	10,291	31,693	5,823	29,298	32,215	109,320

(단위: 명)

- 성별 인력 현황을 살펴보면, 남성 인력이 대부분을 차지하고 있는 것으로 나타났다. 데이터 분석, 데이터 기획·마케터 분야의 경우 여성 비중이 상대적으로 높은 것으로 나타났다.

[표3-13] 전 산업 내 데이터직무 성별 인력 현황

구분		데이터 솔루션	데이터 구축	데이터 컨설팅	데이터 서비스	일반산업	전 산업
DA	남성	564	3,050	426	339	4,354	8,733
	여성	54	342	100	150	692	1,338
	전체	618	3,392	526	489	5,046	10,071
데이터 개발자	남성	3,822	11,419	1,851	7,222	10,202	34,516
	여성	607	2,283	337	1,726	1,785	6,738
	전체	4,429	13,702	2,188	8,948	11,987	41,254
데이터 엔지니어	남성	1,254	5,641	607	3,715	3,173	14,390
	여성	214	888	89	719	334	2,244
	전체	1,468	6,529	696	4,434	3,507	16,634
데이터 분석가	남성	492	880	278	1,397	2,809	5,856
	여성	137	758	57	793	797	2,542
	전체	629	1,638	335	2,190	3,606	8,398
DBA	남성	542	3,383	349	5,711	5,146	15,131
	여성	160	349	39	1,318	866	2,732
	전체	702	3,732	388	7,029	6,012	17,863
데이터 사이언티스트	남성	198	369	104	479	437	1,587
	여성	20	41	69	3	83	216
	전체	218	410	173	482	520	1,803
데이터 컨설턴트	남성	1,228	1,105	1,026	58	601	4,018
	여성	302	271	205	68	140	986
	전체	1,530	1,376	1,231	126	741	5,004
데이터 기획·마케터	남성	507	659	207	3,437	600	5,410
	여성	190	255	79	2,163	196	2,883
	전체	697	914	286	5,600	796	8,293
전체	남성	8,607	26,506	4,848	22,358	27,322	89,641
	여성	1,684	5,187	975	6,940	4,893	19,679
	전체	10,291	31,693	5,823	29,298	32,215	109,320

(단위: 명)

- 일반산업을 포함한 데이터직무 인력 중 빅데이터 인력²⁵⁾은 총 9,955명이며, 전체의 9.1%를 차지하였다.

[표3-14] 전 산업 내 데이터직무 중 빅데이터 관련 인력 현황

구분	전 산업 내 데이터직무				전 산업 내 데이터직무 중 빅데이터 인력 비중(%)
	전체		빅데이터 관련 인력		
	규모(명)	비중(%)	규모(명)	비중(%)	
DA	10,071	9.2%	-	-	-
데이터 개발자	41,254	37.7%	2,908	29.2%	7.0%
데이터 엔지니어	16,634	15.2%	1,679	16.9%	10.1%
데이터 분석가	8,398	7.7%	1,108	11.1%	13.2%
DBA	17,863	16.3%	-	-	-
데이터 사이언티스트	1,803	1.7%	1,803	18.1%	100.0%
데이터 컨설턴트	5,004	4.6%	1,724	17.3%	34.5%
데이터 기획·마케터	8,293	7.6%	733	7.4%	8.8%
전체	109,320	100.0%	9,955	100.0%	9.1%

(단위: 명, %)

※ 빅데이터 관련 인력 : 데이터직무 인력 중 빅데이터 기술을 보유한 인력으로 데이터직무 인력 수에 포함됨

※ DA와 DBA는 조사 제외 대상이며, 데이터 사이언티스트 직무는 100% 빅데이터 관련 인력으로 간주

- 전 산업의 현재 데이터직무 중 빅데이터 인력 비중을 살펴보면, 일반산업에서의 빅데이터 인력 비중이 11.3%로 데이터산업 보다 비교적 높게 나타났으며, 특히 데이터 컨설턴트와 데이터 기획·마케터의 비중이 높게 나타났다.

25) 빅데이터 관련 기술 예시는 다음과 같음

- ① Hadoop, NoSQL, MapReduce, Spark 등 빅데이터를 처리하거나 프로그램을 개발하기 위한 관련 기술 보유
- ② 두 개 이상의 데이터 포맷(예: 정형, 비정형)과 데이터 소스(예: 내부, 외부)를 활용하는 빅데이터 분석 기술 보유
- ③ 실시간으로 스트리밍 되는 데이터 분석 기술 보유
- ④ 빅데이터 관련 제품 및 서비스 기획 능력 보유
- ⑤ 빅데이터 분석을 통한 문제해결, 의사결정 도출 및 제시 능력
- ⑥ 기타 빅데이터 분석, 처리, 개발 등을 위한 관련 기술 보유 등

[표3-15] 전 산업 내 데이터직무 중 빅데이터 관련 인력 비중

구분	전 산업 내 데이터직무 중 빅데이터 관련 인력 비중		
	데이터산업	일반산업	전 산업
데이터 개발자	6.4%	8.7%	7.0%
데이터 엔지니어	8.2%	17.1%	10.1%
데이터 분석가	8.0%	20.2%	13.2%
데이터 사이언티스트	100.0%	100.0%	100.0%
데이터 컨설턴트	32.1%	47.9%	34.5%
데이터 기획·마케터	4.6%	48.9%	8.8%
전체	8.2%	11.3%	9.1%

※ 데이터 사이언티스트 직무는 100% 빅데이터 관련 인력으로 간주

(단위: %)

- 전 산업에서는 빅데이터 관련 데이터 개발자 비중이 가장 높았으며, 이어서 데이터 사이언티스트, 데이터 컨설턴트 순이었다.

[표3-16] 전 산업 내 부문별 데이터직무 중 빅데이터 관련 인력 현황

구분	전 산업 내 데이터직무 중 빅데이터 관련 인력 현황		
	데이터산업	일반산업	전 산업
데이터 개발자	1,866	1,042	2,908
	29.5%	28.7%	29.2%
데이터 엔지니어	1,081	598	1,679
	17.1%	16.5%	16.9%
데이터 분석가	381	727	1,108
	6.0%	20.0%	11.1%
데이터 사이언티스트	1,283	520	1,803
	20.3%	14.3%	18.1%
데이터 컨설턴트	1,369	355	1,724
	21.7%	9.8%	17.3%
데이터 기획·마케터	344	389	733
	5.4%	10.7%	7.4%
전체	6,324	3,631	9,955
	100.0%	100.0%	100.0%

(단위: 명, %)

3.2 인력 수요

1) 데이터산업 인력 수요

- 향후 3년 내 데이터산업 내 데이터직무에 추가로 필요한 인력은 총 4,865명으로 조사되었다. 이 중 데이터 개발자 수요가 52.7%로 가장 높게 나타났다. 이어서 데이터 엔지니어 15.9%, 데이터 분석가 8.1%, DA 7.3%의 순으로 나타났다.

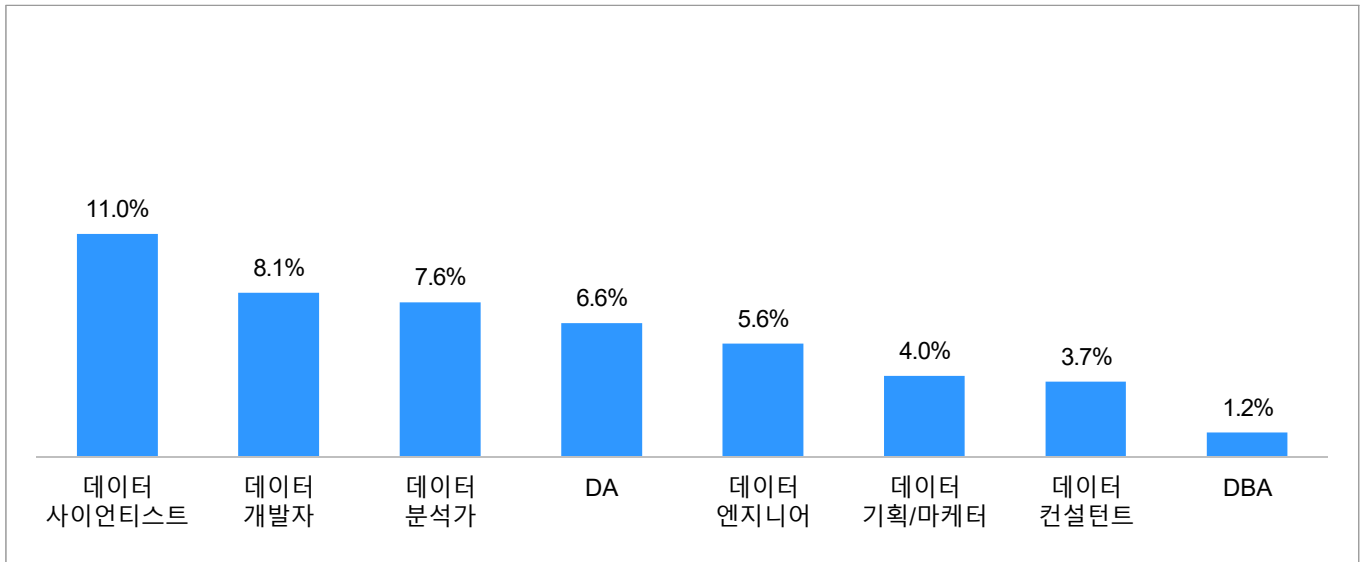
[표3-17] 2017년 데이터산업 내 데이터직무별 필요인력

구분	데이터솔루션	데이터구축	데이터컨설팅	데이터서비스	전체
DA	140	117	44	53	354
	13.5%	6.5%	5.6%	4.3%	7.3%
데이터 개발자	492	1,127	361	584	2,564
	47.4%	62.5%	45.7%	47.2%	52.7%
데이터 엔지니어	163	358	128	126	775
	15.7%	19.9%	16.2%	10.2%	15.9%
데이터 분석가	83	32	127	152	394
	8.0%	1.8%	16.1%	12.3%	8.1%
DBA	15	41	19	70	145
	1.4%	2.3%	2.4%	5.7%	3.0%
데이터 사이언티스트	68	50	29	11	158
	6.6%	2.8%	3.7%	0.9%	3.2%
데이터 컨설턴트	35	19	56	54	164
	3.4%	1.0%	7.1%	4.4%	3.4%
데이터 기획·마케터	42	58	25	186	311
	4.0%	3.2%	3.2%	15.0%	6.4%
전체	1,038	1,802	789	1,236	4,865
	100.0%	100.0%	100.0%	100.0%	100.0%

(단위: 명, %)

- 2017년 데이터직무 인력 부족률을 보면 데이터 사이언티스트(11.0%), 데이터개발자(8.1%), 데이터분석가(7.6%), DA(6.6%)가 전체 평균 부족률 5.9%보다 높은 것으로 나타났다.

[그림3-5] 2017년 데이터산업 내 데이터직무 인력 부족률



[표3-18] 2017년 데이터산업 내 데이터직무 인력 부족률

구분	데이터솔루션	데이터구축	데이터컨설팅	데이터서비스	전체
DA	18.5%	3.3%	7.7%	9.8%	6.6%
데이터 개발자	10.0%	7.6%	14.2%	6.1%	8.1%
데이터 엔지니어	10.0%	5.2%	15.5%	2.8%	5.6%
데이터 분석가	11.7%	1.9%	27.5%	6.5%	7.6%
DBA	2.1%	1.1%	4.7%	1.0%	1.2%
데이터 사이언티스트	23.8%	10.9%	14.4%	2.2%	11.0%
데이터 컨설턴트	2.2%	1.4%	4.4%	30.0%	3.7%
데이터 기획·마케터	5.7%	6.0%	8.0%	3.2%	4.0%
전체	9.2%	5.4%	11.9%	4.0%	5.9%

(단위: %)

※ 인력 부족률 : 필요(부족)인력/(현인력+필요(부족)인력)

- 데이터산업의 데이터직무 수요 중 빅데이터 관련 필요인력은 총 1,380명이며, 전체 필요인력의 28.4%를 차지하였다. 특히 데이터 사이언티스트, 데이터 컨설턴트, 데이터 분석가 순으로 차지하는 비중이 높게 나타났다.

[표3-19] 데이터산업 내 데이터직무 중 빅데이터 관련 필요인력

구분	데이터산업 내 데이터직무 필요인력				데이터산업 내 데이터직무 필요인력 중 빅데이터 필요인력 비중 (%)
	전체		빅데이터 관련 필요인력		
	규모(명)	비중(%)	규모(명)	비중(%)	
DA	354	7.3%	-	-	-
데이터 개발자	2,564	52.7%	660	47.8%	25.7%
데이터 엔지니어	775	15.9%	123	8.9%	15.9%
데이터 분석가	394	8.1%	227	16.5%	57.6%
DBA	145	3.0%	-	-	-
데이터 사이언티스트	158	3.2%	158	11.5%	100.0%
데이터 컨설턴트	164	3.4%	104	7.5%	63.4%
데이터 기획·마케터	311	6.4%	108	7.8%	34.7%
전체	4,865	100.0%	1,380	100.0%	28.4%

(단위: 명, %)

※ 빅데이터 관련 인력 : 데이터직무 인력 중 빅데이터 기술을 보유한 인력으로 데이터직무 인력 수에 포함됨

※ DA와 DBA는 조사 제외 대상이며, 데이터 사이언티스트 직무는 100% 빅데이터 관련 인력으로 간주

- 데이터서비스 분야에서 빅데이터 관련 데이터 개발자 필요인력이 79.7%로 가장 높았으며, 데이터컨설팅 분야는 빅데이터 관련 데이터 분석가, 데이터구축 분야는 빅데이터 관련 데이터 개발자가 높게 나타났다.

[표3-20] 데이터산업 내 부문별 빅데이터 관련 필요인력

구분	데이터산업 내 부문별 빅데이터 관련 필요인력				
	데이터솔루션	데이터구축	데이터컨설팅	데이터서비스	전체
데이터 개발자	151	124	13	372	660
	37.7%	41.5%	6.1%	79.7%	47.8%
데이터 엔지니어	31	62	8	22	123
	7.7%	20.7%	3.8%	4.7%	8.9%
데이터 분석가	79	17	117	14	227
	19.7%	5.7%	54.9%	3.0%	16.5%
데이터 사이언티스트	68	50	29	11	158
	17.0%	16.7%	13.6%	2.4%	11.5%
데이터 컨설턴트	17	18	38	31	104
	4.2%	6.0%	17.8%	6.6%	7.5%
데이터 기획·마케터	55	28	8	17	108
	13.7%	9.4%	3.8%	3.6%	7.8%
전체	401	299	213	467	1,380
	100.0%	100.0%	100.0%	100.0%	100.0%

(단위: 명, %)

○ 2017년 현재 빅데이터 관련 인력 부족률을 보면 데이터 분석가가 가장 높고, 다음으로 데이터 개발자, 데이터 기획/마케터의 순으로 높게 나타났다.

[표3-21] 2017년 데이터산업 내 데이터직무 중 빅데이터 관련 인력 부족률

구분	데이터산업 내 데이터직무 중 빅데이터 관련 인력 부족률				
	데이터솔루션	데이터구축	데이터컨설팅	데이터서비스	전체
데이터 개발자	14.4%	23.3%	3.6%	63.8%	26.1%
데이터 엔지니어	3.6%	31.2%	12.3%	27.5%	10.2%
데이터 분석가	68.1%	11.3%	37.4%	50.0%	37.3%
데이터 사이언티스트	23.8%	10.9%	14.4%	2.2%	11.0%
데이터 컨설턴트	5.9%	3.6%	6.3%	37.3%	7.1%
데이터 기획·마케터	17.2%	51.9%	33.3%	30.9%	23.9%
전체	13.8%	15.7%	13.6%	35.3%	17.9%

※ 인력 부족률 : 필요(부족)인력/(현인력+필요(부족)인력)

(단위: %)

2) 전 산업 데이터직무 인력 수요

- 향후 3년 내, 일반산업을 포함한 전 산업에서 추가로 필요로 하는 데이터직무 인력은 총 1만 3,337명으로 조사되었다. 이 중 데이터 개발자 수요가 37.2%로 가장 높게 나타났다.
- 데이터 개발자 이외의 필요인력을 살펴보면, 데이터산업은 데이터 엔지니어가, 일반산업은 DA와 DBA의 수요가 상대적으로 높게 나타났다.

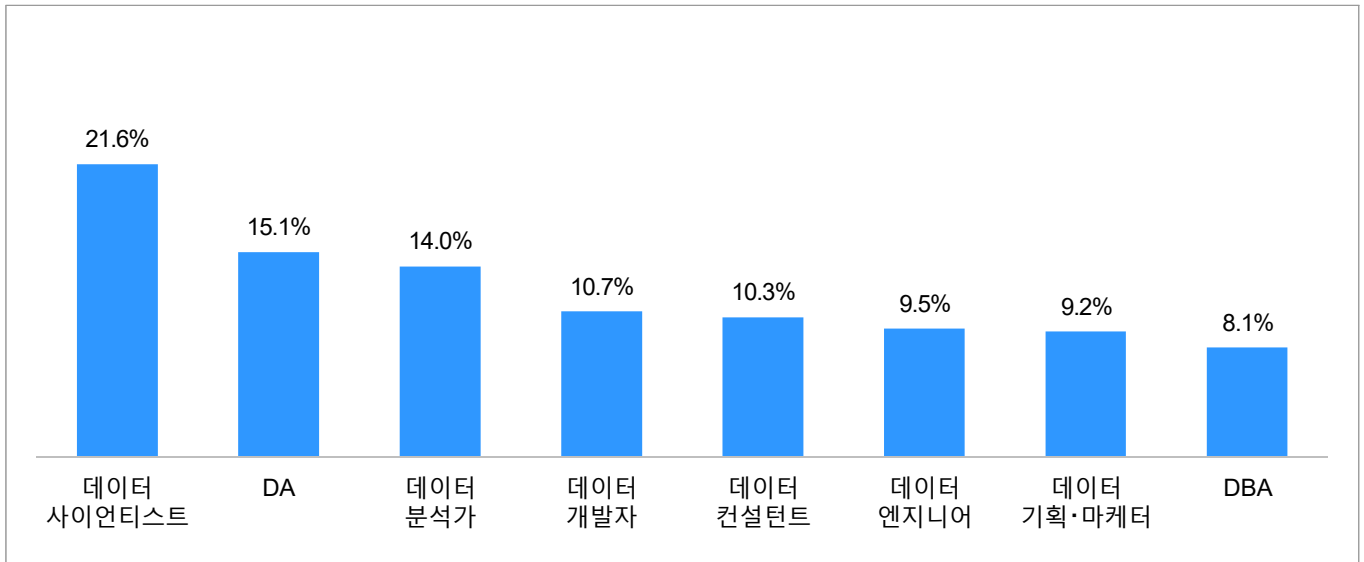
[표3-22] 향후 3년 내 전 산업 내 데이터직무별 필요인력

구분	데이터산업	일반산업	전 산업
DA	354	1,437	1,791
	7.3%	17.0%	13.4%
데이터 개발자	2,564	2,389	4,953
	52.7%	28.2%	37.2%
데이터 엔지니어	775	963	1,738
	15.9%	11.4%	13.0%
데이터 분석가	394	977	1,371
	8.1%	11.5%	10.3%
DBA	145	1,424	1,569
	3.0%	16.8%	11.8%
데이터 사이언티스트	158	338	496
	3.2%	4.0%	3.7%
데이터 컨설턴트	164	410	574
	3.4%	4.8%	4.3%
데이터 기획·마케터	311	534	845
	6.4%	6.3%	6.3%
전체	4,865	8,472	13,337
	100.0%	100.0%	100.0%

(단위: 명, %)

- 전체적으로는 데이터 개발자에 대한 수요가 높은 반면, 2017년 데이터직무 인력 부족률은 데이터 사이언티스트가 가장 높고, 다음으로 DA, 데이터 분석가 등의 순으로 나타났다.

[그림3-6] 2017년 전 산업 내 데이터직무별 인력 부족률



[표3-23] 2017년 전 산업 내 데이터직무별 인력 부족률

구분	데이터산업	일반산업	전 산업
DA	6.6%	22.2%	15.1%
데이터 개발자	8.1%	16.6%	10.7%
데이터 엔지니어	5.6%	21.5%	9.5%
데이터 분석가	7.6%	21.3%	14.0%
DBA	1.2%	19.2%	8.1%
데이터 사이언티스트	11.0%	39.4%	21.6%
데이터 컨설턴트	3.7%	35.6%	10.3%
데이터 기획·마케터	4.0%	40.2%	9.2%
전체	5.9%	20.8%	10.9%

※ 인력 부족률 : 필요(부족)인력/(현인력+필요(부족)인력)

(단위: %)

- 일반산업을 포함한 전 산업의 데이터직무 필요인력 중 빅데이터 필요인력은 총 6,008명이며, 전체 수요의 45.0%를 차지하고 있다.

[표3-24] 전 산업 내 데이터직무 중 빅데이터 관련 필요인력

구분	전 산업 내 데이터직무 필요인력				전 산업 내 데이터직무 필요인력 중 빅데이터 관련 필요인력 비중 (%)
	전체		빅데이터 관련 필요인력		
	규모(명)	비중(%)	규모(명)	비중(%)	
DA	1,791	13.4%	-	-	-
데이터 개발자	4,953	37.2%	2,350	39.1%	47.4%
데이터 엔지니어	1,738	13.0%	907	15.1%	52.2%
데이터 분석가	1,371	10.3%	1,331	22.1%	97.1%
DBA	1,569	11.8%	-	-	-
데이터 사이언티스트	496	3.7%	496	8.3%	100.0%
데이터 컨설턴트	574	4.3%	509	8.5%	88.7%
데이터 기획·마케터	845	6.3%	415	6.9%	49.1%
전체	13,337	100.0%	6,008	100.0%	45.0%

(단위: 명, %)

※ 빅데이터 관련 인력 : 데이터직무 인력 중 빅데이터 기술을 보유한 인력으로 데이터직무 인력 수에 포함됨

※ DA와 DBA는 조사 제외 대상이며, 데이터 사이언티스트 직무는 100% 빅데이터 관련 인력으로 간주

- 빅데이터 관련 필요인력은 데이터 개발자가 39.1%로 가장 수요가 높고, 다음으로 데이터 분석가(22.1%), 데이터 엔지니어(15.1%) 등의 순으로 나타났다.

[표3-25] 전 산업 내 부문별 데이터직무 중 빅데이터 관련 필요인력

구분	전 산업 내 데이터직무 중 빅데이터 필요인력		
	데이터산업	일반산업	전 산업
데이터 개발자	660	1,690	2,350
	47.8%	36.5%	39.1%
데이터 엔지니어	123	784	907
	8.9%	16.9%	15.1%
데이터 분석가	227	1,104	1,331
	16.5%	23.9%	22.1%
데이터 사이언티스트	158	338	496
	11.5%	7.3%	8.3%
데이터 컨설턴트	104	405	509
	7.5%	8.8%	8.5%
데이터 기획·마케터	108	307	415
	7.8%	6.6%	6.9%
전체	1,380	4,628	6,008
	100.0%	100.0%	100.0%

(단위: 명, %)

- 전 산업의 데이터직무 중 빅데이터 인력 부족률을 보면 빅데이터 분석가가 가장 높고, 다음으로 빅데이터 개발자, 빅데이터 기획·마케터, 빅데이터 엔지니어 등의 순으로 나타났다.

[표3-26] 2017년 전 산업 내 데이터직무 중 빅데이터 인력 부족률

구분	전 산업 내 데이터직무 중 빅데이터 인력 부족률		
	데이터산업	일반산업	전 산업
데이터 개발자	26.1%	61.9%	44.7%
데이터 엔지니어	10.2%	56.7%	35.1%
데이터 분석가	37.3%	60.3%	54.6%
데이터 사이언티스트	11.0%	39.4%	21.6%
데이터 컨설턴트	7.1%	53.3%	22.8%
데이터 기획·마케터	23.9%	44.1%	36.1%
전체	17.9%	56.0%	37.6%

※ 인력 부족률 : 필요(부족)인력/(현인력+필요(부족)인력)

(단위: %)

3.3 데이터 자격증

- 데이터산업에서 가장 우대하는 데이터 관련 전문 인력 자격증은 정보처리기사(26.3%)가 가장 많은 것으로 나타났다.

[표3-27] 데이터산업의 데이터 전문 인력 우대 자격증

구분	응답수	데이터 아키텍처 전문가 (DAP)	데이터 아키텍처 준 전문가 (DAsP)	SQL 전문가 (SQLP)	SQL 개발자 (SQLD)	데이터 분석 전문가 (ADP)	데이터 분석 준 전문가 (ADsP)	정보 처리 기사	사회 조사 분석사 1급	사회 조사 분석사 2급	기타	없다
데이터산업 전체	(819)	2.9	1.1	1.8	2.8	2.2	0.5	26.3	0.1	0.2	1.8	67.5
데이터솔루션	(231)	5.2	1.7	3.0	3.9	2.2	0.4	31.2	0.4	0.4	1.7	58.9
데이터구축	(233)	1.7	1.3	2.1	3.4	1.7	0.0	40.8	0.0	0.0	1.7	56.2
데이터컨설팅	(77)	5.2	1.3	1.3	1.3	1.3	1.3	32.5	0.0	0.0	6.5	55.8
데이터서비스	(278)	1.4	0.4	0.7	1.8	2.9	0.7	8.3	0.0	0.4	0.7	87.4

※ 복수응답

(단위: 개, %)

- 데이터산업에서는 최근 정보처리기사의 비중이 가장 높은 것으로 나타났다.

[표3-28] 데이터산업의 연도별 우대 자격증 비교

구분	응답수	데이터 아키텍처 전문가 (DAP)	데이터 아키텍처 준 전문가 (DAsP)	SQL 전문가 (SQLP)	SQL 개발자 (SQLD)	데이터 분석 전문가 (ADP)	데이터 분석 준 전문가 (ADsP)	정보 처리 기사	사회 조사 분석사 1급	사회 조사 분석사 2급	기타	없다
2017년	(819)	2.9	1.1	1.8	2.8	2.2	0.5	26.3	0.1	0.2	1.8	67.5
2016년	(730)	10.3	8.2	15.3	22.9	17.7	4.1	31.2	0.8	-	3.2	32.2
2015년	(521)	7.2	8.5	20.5	13.3	6.4	2.9	4.1	-	-	2.3	60.2

※ 복수응답

(단위: 개, %)

- 일반산업을 포함해 전 산업에서 우대하는 데이터 관련 전문 인력 자격증은 정보처리기사(28.5%), SQL 전문가(12.4%), SQL 개발자(12.1%) 순으로 나타났다.

[표3-29] 전 산업의 데이터 전문 인력 우대 자격증

구분	응답수	데이터 아키텍처 전문가 (DAP)	데이터 아키텍처 준 전문가 (DAsP)	SQL 전문가 (SQLP)	SQL 개발자 (SQLD)	데이터 분석 전문가 (ADP)	데이터 분석 준 전문가 (ADsP)	정보 처리 기사	사회 조사 분석사 1급	사회 조사 분석사 2급	기타	없다
데이터산업	(819)	2.9	1.1	1.8	2.8	2.2	0.5	26.3	0.1	0.2	1.8	67.5
일반산업	(875)	10.1	5.1	22.3	20.8	7.0	2.9	30.5	1.5	0.0	4.0	39.9
전 산업	(1,694)	6.6	3.2	12.4	12.1	4.7	1.7	28.5	0.8	0.1	3.0	53.2

※ 복수응답

(단위: 개, %)

- 자격증별 우대사항이 동일한 경우, 데이터산업에 종사하는 기업들이 데이터 관련 자격증과 관련해 제공하는 우대사항으로는 입사우대가 82.4%로 가장 높았으며, 그 밖에 인사고과 반영, 수당 지급 등의 형태가 뒤를 잇고 있다.
- 자격증별 우대사항이 동일한 경우, 일반산업이 포함된 전 산업의 데이터 관련 자격증에 대한 우대사항으로 입사 우대가 가장 높았으며, 그 다음은 인사고과 반영이 높게 나타났다.

[표3-30] 전 산업의 데이터 전문 인력 우대사항

구분	응답수	입사 우대	응시료 지원	수당 지급	인사고과 반영	기타	없다
데이터산업	(17)	82.4	-	5.9	17.6	0.0	5.9
일반산업	(416)	36.8	2.2	8.2	8.7	0.0	50.7
전 산업	(433)	38.6	2.1	8.1	9.0	0.0	49.0

※ 복수응답, 자격증별 우대사항이 동일한 경우

(단위: 개, %)

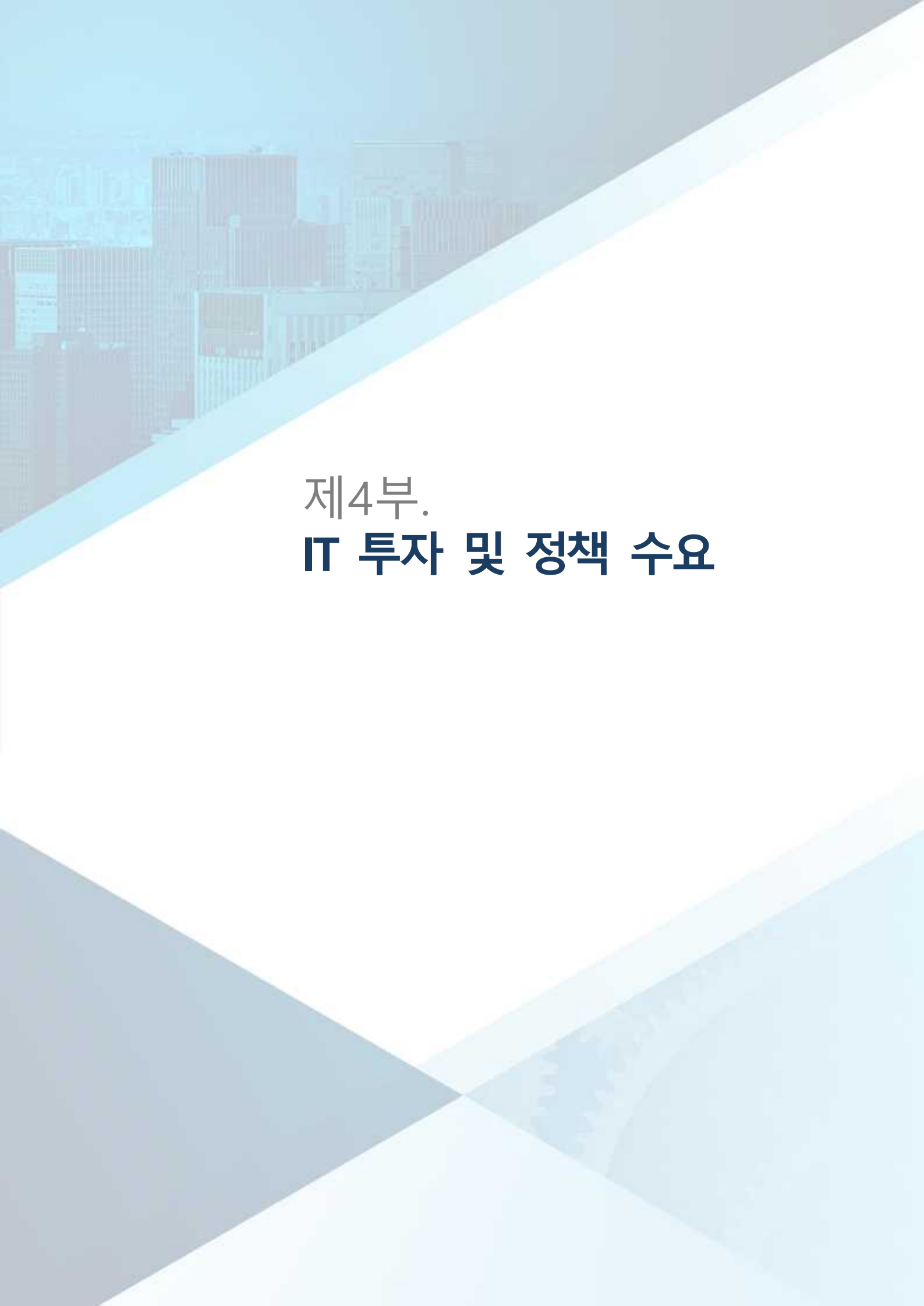
- 자격증별 우대사항이 다를 경우, 전 산업의 데이터 관련 자격증별 우대사항을 보면, 대체로 입사 우대를 제공하고 있으며, 전반적으로 정보처리기사가 입사우대를 받는 것으로 나타났다.
- 데이터서비스 기업의 경우 보유 자격증에 따라 인사고과에 반영하는 비율이 타 분야 대비 상대적으로 높게 나타났다.

[표3-31] 전 산업의 자격증별 데이터 전문 인력 우대사항

구분		응답수	입사 우대	응시료 지원	수당 지급	인사고과 반영	없다
데이터 솔루션	데이터 아키텍처 전문가(DAP)	(88)	12.5	0.0	0.0	0.0	87.5
	데이터 아키텍처 준전문가(DAsP)	(88)	4.5	0.0	0.0	0.0	95.5
	SQL 전문가(SQLP)	(88)	6.8	0.0	0.0	1.1	92.0
	SQL 개발자(SQLD)	(90)	8.9	0.0	0.0	2.2	88.9
	데이터 분석 전문가(ADP)	(89)	4.5	0.0	1.1	1.1	93.3
	데이터 분석 준전문가(ADsP)	(88)	1.1	0.0	0.0	0.0	98.9
	정보처리기사	(94)	64.9	0.0	3.2	10.6	21.3
	기타	(89)	5.6	0.0	0.0	0.0	94.4
데이터 구축	데이터 아키텍처 전문가(DAP)	(98)	3.1	0.0	0.0	1.0	95.9
	데이터 아키텍처 준전문가(DAsP)	(98)	3.1	0.0	0.0	1.0	95.9
	SQL 전문가(SQLP)	(97)	3.1	0.0	0.0	0.0	96.9
	SQL 개발자(SQLD)	(97)	3.1	0.0	0.0	1.0	95.9
	데이터 분석 전문가(ADP)	(97)	2.1	0.0	0.0	0.0	97.9
	데이터 분석 준전문가(ADsP)	(98)	1.0	0.0	0.0	0.0	99.0
	정보처리기사	(102)	85.3	1.0	3.9	4.9	4.9
	기타	(97)	3.1	0.0	0.0	1.0	95.9
데이터 컨설팅	데이터 아키텍처 전문가(DAP)	(34)	8.8	2.9	2.9	0.0	85.3
	데이터 아키텍처 준전문가(DAsP)	(33)	3.0	0.0	0.0	0.0	97.0
	SQL 전문가(SQLP)	(34)	2.9	0.0	2.9	0.0	94.1
	SQL 개발자(SQLD)	(33)	3.0	0.0	0.0	0.0	97.0
	데이터 분석 전문가(ADP)	(34)	2.9	0.0	2.9	0.0	94.1
	데이터 분석 준전문가(ADsP)	(35)	8.6	0.0	0.0	0.0	91.4
	정보처리기사	(35)	68.6	0.0	5.7	0.0	25.7
	기타	(33)	12.1	0.0	0.0	3.0	84.8
데이터 서비스	데이터 아키텍처 전문가(DAP)	(32)	6.3	0.0	0.0	6.3	87.5
	데이터 아키텍처 준전문가(DAsP)	(31)	3.2	0.0	0.0	3.2	93.5
	SQL 전문가(SQLP)	(31)	3.2	0.0	0.0	3.2	93.5
	SQL 개발자(SQLD)	(32)	12.5	0.0	0.0	6.3	81.3
	데이터 분석 전문가(ADP)	(31)	16.1	0.0	0.0	3.2	80.6
	데이터 분석 준전문가(ADsP)	(28)	0.0	0.0	0.0	0.0	100.0
	정보처리기사	(31)	64.5	0.0	0.0	3.2	32.3
	기타	(34)	5.9	0.0	0.0	0.0	94.1
일반 산업	데이터 아키텍처 전문가(DAP)	(32)	87.5	9.4	6.3	15.6	0.0
	데이터 아키텍처 준전문가(DAsP)	(15)	80.0	6.7	6.7	20.0	0.0
	SQL 전문가(SQLP)	(61)	90.2	9.8	6.6	9.8	0.0
	SQL 개발자(SQLD)	(49)	81.6	6.1	6.1	12.2	0.0
	데이터 분석 전문가(ADP)	(27)	88.9	11.1	7.4	18.5	0.0
	데이터 분석 준전문가(ADsP)	(15)	80.0	6.7	6.7	26.7	0.0
	정보처리기사	(73)	64.4	5.5	17.8	11.0	0.0
	기타	(8)	75.0	0.0	12.5	37.5	0.0

※ 복수응답, 자격증별 우대사항이 동일한 경우

(단위: 개, %)

The background features a stylized city skyline in shades of blue and grey on the left side. A large, light blue diagonal band runs from the top right towards the bottom left. In the bottom right corner, there are faint, overlapping geometric shapes, including a gear and a circle, in a light blue color.

제4부. **IT 투자 및 정책 수요**

4.1 IT 투자 현황

- 데이터산업은 기업 전체 매출액 중 데이터 관련 사업에 투자하는 비중이 2016년 20.4%이며, 2017년 예상 투자 비중은 20.5%로 비슷한 수준으로 나타났다.
- 데이터솔루션 기업은 2016년과 2017년도 모두 매출액 대비 데이터 관련 사업 투자 비중이 각각 23.5%, 23.2%로 타 분야에 비해 투자 비중이 높은 것으로 나타났다.
- 데이터서비스 기업은 2016년과 2017년도 모두 매출액 대비 투자 비중이 약 17%대로 타 분야에 비해 낮게 나타났다.

[표4-1] 데이터 관련 사업 투자비중

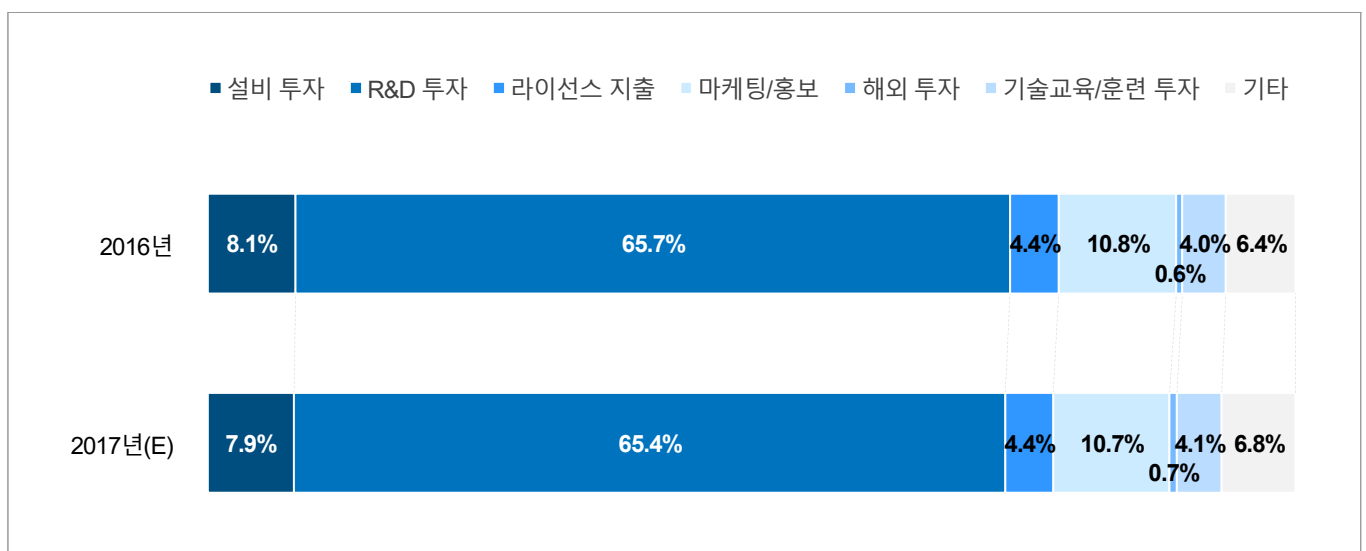
구분	응답수	2016년	2017년(E)
데이터산업 전체	(1004)	20.4	20.5
데이터솔루션	(274)	23.5	23.2
데이터구축	(262)	22.1	22.3
데이터컨설팅	(86)	17.4	18.7
데이터서비스	(382)	17.2	17.4

※ 투자비중 : 매출 대비 투자 비중을 의미함

(단위: 개, %)

- 데이터산업 전체의 2017년 세부 투자 비중을 살펴보면, R&D 투자 비중이 65.4%로 가장 높았고, 다음으로 마케팅 및 홍보, 설비투자, 라이선스 지출, 기술교육 및 훈련투자 등의 순으로 나타났다.

[그림4-1] 데이터산업 전체 세부 투자비중



- R&D투자 비중은 데이터구축, 솔루션, 컨설팅, 서비스 기업 순으로 높게 나타났다. 데이터서비스 기업은 2016년과 2017년 모두 마케팅 및 홍보에 대한 투자 비중이 타 분야 대비 상대적으로 높게 나타났다.

[표4-2] 데이터산업 분야별 세부 투자비중

구분		응답수	설비 투자	R&D 투자	라이선 스 지출	마케팅/ 홍보	해외 투자	기술 교육/ 훈련 투자	기타
2016년	데이터산업 전체	(1004)	8.1	65.7	4.4	10.8	0.6	4.0	6.4
	데이터솔루션	(274)	6.6	71.4	5.1	6.4	0.6	6.1	3.8
	데이터구축	(262)	8.4	73.6	2.7	8.5	0.3	3.0	3.5
	데이터컨설팅	(86)	8.1	67.8	4.2	6.2	0.0	6.5	7.2
	데이터서비스	(382)	9.0	54.1	5.3	17.7	0.8	2.3	10.8
2017년(E)	데이터산업 전체	(1004)	7.9	65.4	4.4	10.7	0.7	4.1	6.8
	데이터솔루션	(274)	6.4	71.3	5.0	6.3	1.1	6.5	3.4
	데이터구축	(262)	8.3	73.7	2.7	8.3	0.3	3.0	3.7
	데이터컨설팅	(86)	8.1	63.6	4.2	6.2	0.0	6.5	11.4
	데이터서비스	(382)	8.9	54.5	5.2	17.5	0.8	2.3	10.8

(단위: 개, %)

4.2 시장 환경 현황

- 데이터 관련 주력 사업 분야에서 국내기업을 주 고객으로 둔 기업이 58.1%로 절반 이상을 차지하고 있으며, 정부 및 공공기관을 대상으로 하는 기업은 18.9%로 나타났다.
- 데이터솔루션과 컨설팅 기업의 경우 국내 기업 비중이 전체의 2/3 이상을 차지하는 것으로 나타났으며, 데이터서비스 기업에서는 국내 개인이 주 고객인 기업이 42.4%로 타 분야 대비 상대적으로 높게 나타났다.

[표4-3] 데이터 관련 사업 주요 고객

구분	응답수	국내			해외			기타	무응답
		기업	개인	정부/공공	기업	개인	정부/공공		
데이터산업 전체	(1004)	58.1	18.3	18.9	0.1	0.1	0.1	0.1	4.3
데이터솔루션	(274)	70.4	4.4	20.1	0.0	0.0	0.4	0.4	4.3
데이터구축	(262)	58.0	2.7	35.1	0.4	0.4	0.0	0.0	3.4
데이터컨설팅	(86)	68.6	3.5	25.6	0.0	0.0	0.0	0.0	2.3
데이터서비스	(382)	46.9	42.4	5.5	0.0	0.0	0.0	0.0	5.2

(단위: 개, %)

- 데이터 관련 주력 사업 분야의 주요 경쟁국으로 70.0%가 한국을 꼽고 있는 것으로 나타났다. 그 외 미국, 중국, 일본, 유럽 등의 순이며, 기타로는 대만 등으로 나타났다.

[표4-4] 데이터 관련 사업 주요 경쟁국

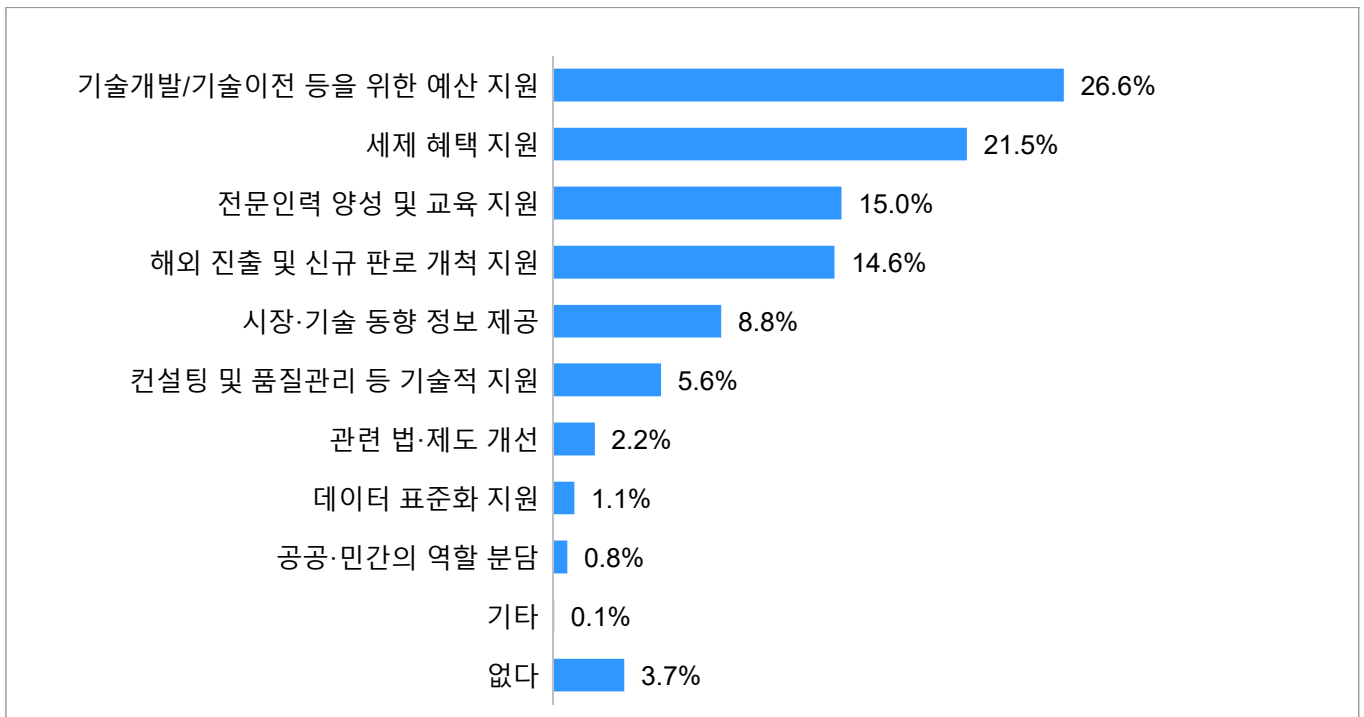
구분	응답수	미국	중국	유럽	일본	한국	기타	무응답
데이터산업 전체	(1004)	6.2	3.2	0.9	1.0	70.0	14.7	4.0
데이터솔루션	(274)	11.3	4.7	1.5	1.8	66.1	10.6	4.0
데이터구축	(262)	5.3	3.1	0.4	0.8	72.1	14.9	3.4
데이터컨설팅	(86)	5.8	1.2	0.0	0.0	83.7	7.0	2.3
데이터서비스	(382)	3.1	2.6	1.0	0.8	68.3	19.4	4.8

(단위: 개, %)

4.3 정책 수요

- 데이터 관련 사업 추진 및 산업 활성화를 위해 가장 필요한 정책 지원 사항으로 기술개발/기술 이전 등을 위한 예산 지원이 26.6%로 가장 높게 나타났다.
- 다음으로 세제혜택 지원, 전문인력 양성 및 교육 지원, 해외 진출 및 신규 판로 개척 지원 등에 대한 수요 순으로 나타났다.

[그림4-2] 데이터산업 정책 수요



- 데이터솔루션 기업은 전문인력 양성 및 교육 지원에 대한 수요가 상대적으로 높았으며, 데이터 서비스 기업은 시장·기술 동향 정보 제공에 대한 수요가 상대적으로 높게 나타났다.

[표4-5] 데이터산업 부문별 정책 수요

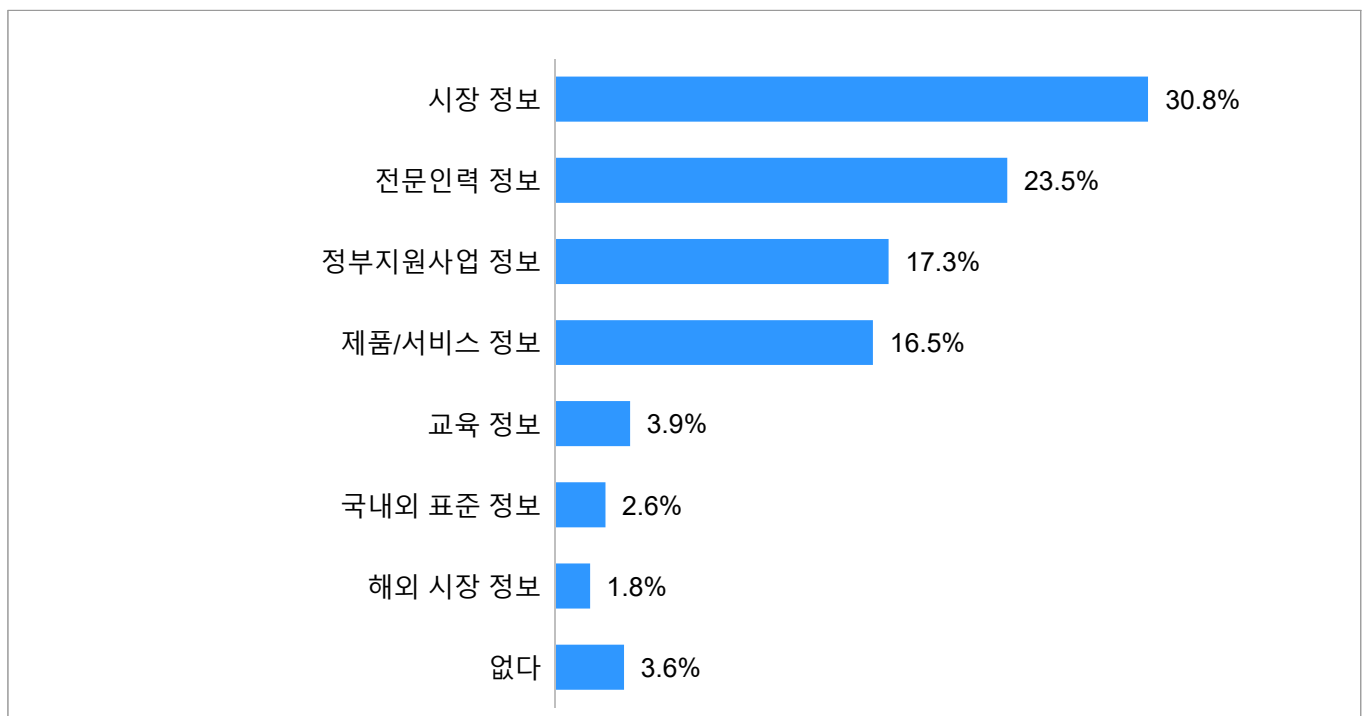
구분	응답수	기술개발/기술이전 등을 위한 예산(자금) 지원	세제 혜택 지원	해외 진출 및 신규 판로 개척 지원	전문인력 양성 및 교육 지원	시장·기술 동향 정보 제공	컨설팅 및 품질 관리 등 기술적 지원	관련 법·제도 개선	데이터 표준화 지원	공공·민간의 역할 분담	기타	없다
데이터산업 전체	(1004)	26.6	21.5	14.6	15.0	8.8	5.6	2.2	1.1	0.8	0.1	3.7
데이터솔루션	(274)	29.3	15.4	15.2	19.3	7.1	6.8	1.5	1.3	1.1	0.0	3.0
데이터구축	(262)	25.0	24.2	16.8	15.5	7.1	4.6	2.0	0.4	0.7	0.3	3.4
데이터컨설팅	(86)	26.6	23.2	11.8	16.0	6.3	6.3	4.6	0.8	0.8	0.0	3.6
데이터서비스	(382)	25.6	23.6	13.4	11.3	11.6	5.3	2.3	1.5	0.6	0.0	4.8

※ 우선순위 가중치 적용

(단위: 개, %)

- 데이터 관련 사업 수행을 위한 필요 정보로는 데이터산업 전반의 시장동향 등 시장정보가 필요하다는 의견이 30.8%로 가장 높게 나타났다. 다음으로 전문인력 정보, 정부지원사업 정보, 제품/서비스 정보 등의 순으로 나타났다.

[그림4-3] 데이터산업 필요 정보



- 데이터솔루션 기업은 전문인력 정보에 대한 필요성이 상대적으로 높았으며, 데이터컨설팅 기업은 정부지원사업 정보, 데이터서비스 기업은 제품·서비스 정보에 대한 수요가 비교적 높게 나타났다.

[표4-6] 데이터산업 부문별 필요 정보

구분	응답수	시장 정보	전문인력 정보	정부지원 사업 정보	제품/서비스 정보	교육 정보	국내외 표준 정보	해외 시장 정보	없다
데이터산업 전체	(1004)	30.8	23.5	17.3	16.5	3.9	2.6	1.8	3.6
데이터솔루션	(274)	30.4	27.6	16.6	14.8	3.0	2.9	1.8	2.9
데이터구축	(262)	33.1	22.8	17.7	14.7	3.7	3.3	1.6	3.1
데이터컨설팅	(86)	34.8	23.5	21.3	9.1	3.0	3.5	2.2	2.6
데이터서비스	(382)	28.7	21.0	16.8	20.7	5.0	1.8	1.9	4.1

※ 우선순위 가중치 적용

(단위: 개, %)

- 데이터 관련 사업 수행의 애로사항을 살펴보면, 동종업계 내 경쟁 심화(25.2%)가 가장 높게 나타났다. 다음으로 고정비용 상승, 시장 포화, 기술개발 부담 등 자금난, 데이터 전문 인력 부족 등으로 나타났다.

[표4-7] 데이터사업 수행 시 애로사항

구분	데이터산업 전체	데이터 솔루션	데이터 구축	데이터 컨설팅	데이터 서비스
응답수	(1004)	(274)	(262)	(86)	(382)
동종업계 내 경쟁심화	25.2	25.8	27.2	26.1	23.2
고정비용(임대료/ 인건비 등) 상승	17.7	14.5	18.1	19.9	19.2
시장 포화	15.6	14.9	13.1	7.5	19.6
기술개발 부담 등 자금난	12.1	14.2	10.1	14.6	11.4
데이터 전문 인력 부족	11.4	16.6	12.2	10.2	7.3
내수시장 악화	8.0	6.3	9.5	12.4	7.1
수익모델 부재	2.0	1.7	0.7	1.3	3.4
신규 기술개발(R&D) 역량 미흡	1.5	0.8	2.6	1.3	1.2
유료화인식부족	1.2	0.6	0.9	0.9	1.9
국제경기 침체	0.7	0.4	1.1	0.9	0.5
데이터 표준화 미비	0.5	0.6	0.3	0.0	0.7
법·제도적 제약	0.5	0.3	0.4	2.2	0.3
해외 판로 개척 어려움	0.2	0.4	0.1	0.0	0.2
비데이터 기업의 사업추진	0.1	0.0	0.3	0.0	0.0
없다	3.3	2.9	3.4	2.7	4.0

※ 우선순위 가중치 적용

(단위: 개, %)

- 한편, 데이터서비스 사업을 하는 기업들은 기술 관련 자산²⁶⁾이 전체 매출에 기여하는 비중이 38.8%인 것으로 나타났으며, 그 중 데이터베이스가 전체 매출에 기여하는 비중은 35.4%로 나타났다.

[표4-8] 기술관련 자산 및 데이터베이스의 매출 기여 비중

구분	기술관련 자산이 전체 매출에 기여하는 비중		데이터베이스가 전체 매출에 기여하는 비중	
	응답수	비중	응답수	비중
데이터서비스 전체	(382)	38.8	(382)	35.4
데이터 거래	(53)	47.2	(53)	44.3
정보 제공	(306)	36.4	(306)	32.9
데이터 분석 제공	(23)	48.5	(23)	46.1

(단위: 개, %)

26) 국제회계기준 상의 5가지 무형자산 유형 중 데이터베이스, 특허기술 등을 의미함



제5부. 데이터 유통 현황

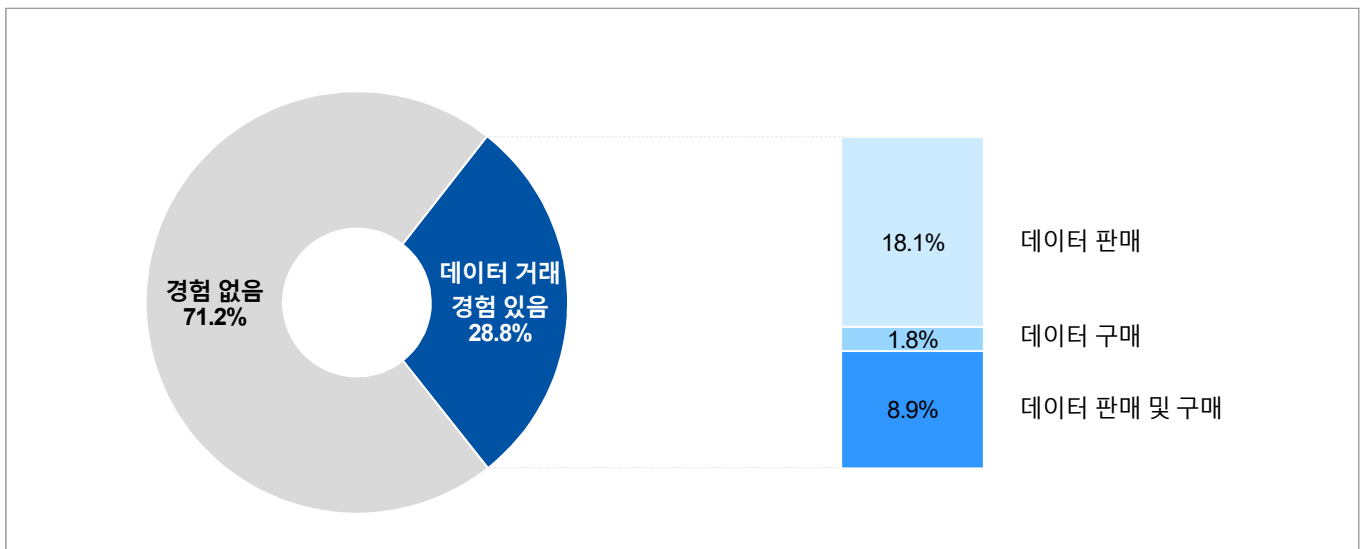
※ 데이터서비스 기업 대상 분석 결과

5.1 데이터 유통 현황

1) 데이터 거래 현황

- 데이터서비스 기업을 대상으로 데이터 유통 현황을 조사한 결과, 데이터서비스 조사기업의 28.8%가 데이터거래 경험이 있는 것으로 나타났다. 그 중 데이터 판매 비중은 18.1%, 데이터 구매 비중은 1.8%, 데이터 판매와 구매를 함께 한 기업은 8.9%로 나타났다.

[그림5-1] 데이터 거래 경험 여부



- 데이터서비스 세부 분야별로 보면, 데이터 거래 기업은 데이터를 판매한 기업이 73.6%로 대부분을 차지한 것으로 나타났다.

[표5-1] 데이터서비스 분야별 데이터 거래 경험 여부

구분	응답수	데이터 판매	데이터 구매	데이터 판매 및 구매	경험 없음
데이터서비스 전체	(382)	18.1	1.8	8.9	71.2
데이터 거래	(53)	73.6	1.9	24.5	0.0
정보 제공	(306)	9.8	1.3	6.2	82.7
데이터 분석 제공	(23)	0.0	8.7	8.7	82.6

(단위: 개, %)

- 데이터 거래 경험이 없는 기업의 경우, 필요한 데이터를 자체적으로 구축한다는 의견이 비교적 높게 나타났다.

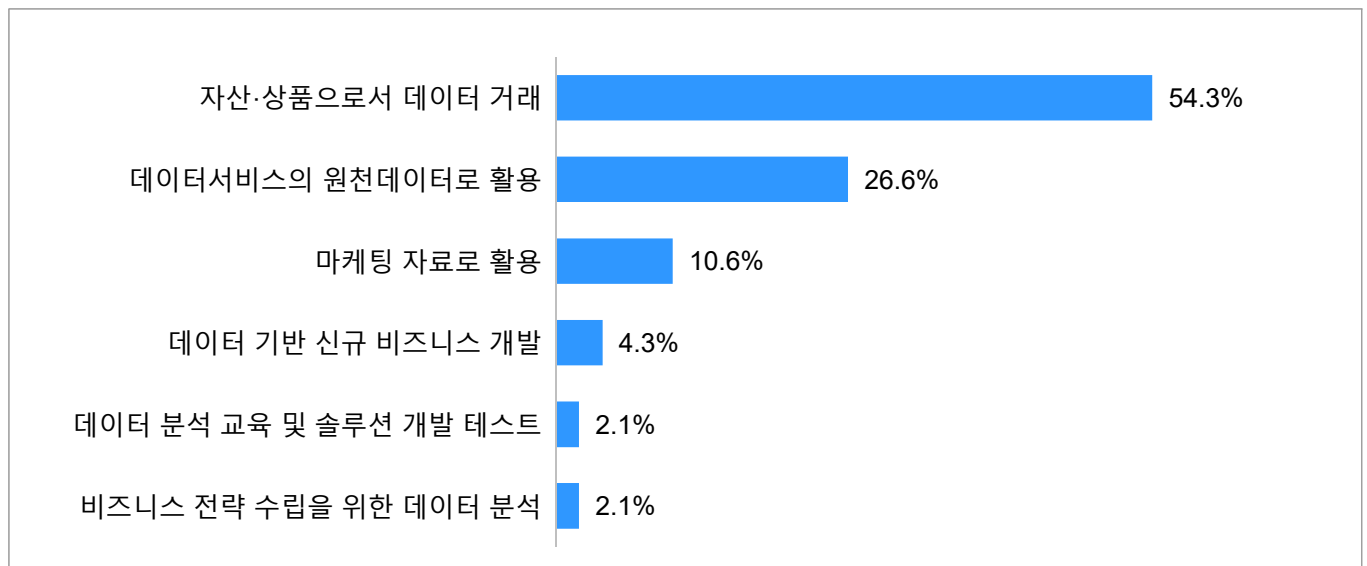
[표5-2] 데이터 거래 경험 없는 이유

구분	응답수	데이터 유통 채널 부재	내부 인력 부족	필요한 데이터 자체 구축	쓸만한 양질의 데이터 부재	사업의 특성상 거래 불필요
데이터서비스 전체	(272)	2.2	0.7	18.4	1.5	77.2
정보 제공	(253)	2.4	0.4	17.4	1.2	78.6
데이터 분석 제공	(19)	0.0	5.3	31.6	5.3	57.8

(단위: 개, %)

- 데이터 거래 경험이 있는 기업의 데이터 거래 목적은 자산 및 상품으로 데이터를 거래한다는 의견이 가장 많이 차지했다.
- 다음으로 데이터서비스의 원천 데이터로 활용, 마케팅 자료로 활용, 데이터 기반 신규 비즈니스 개발 등의 순으로 나타났다.

[그림5-2] 데이터 거래 목적

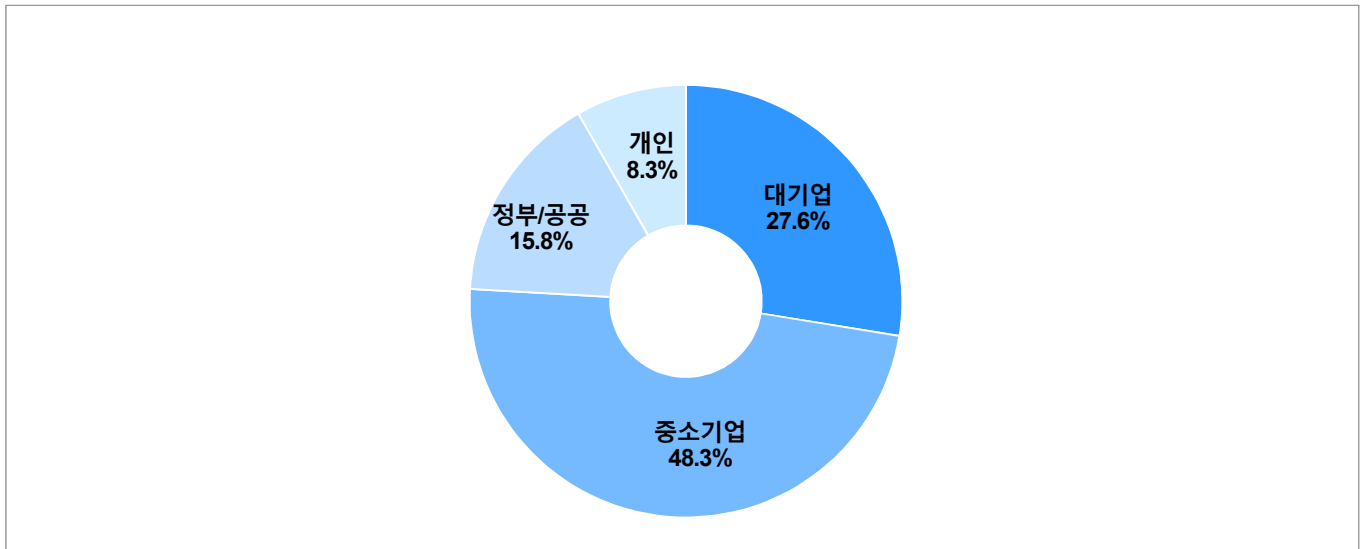


※ 우선순위 가중치 적용

(단위: 개, %)

- 데이터 거래가 있는 기업의 48.3%는 중소기업, 27.6%는 대기업에서 데이터를 구매하는 것으로 나타났고, 다음으로 정부 및 공공분야는 15.9%, 개인은 8.3%로 나타났다.

[그림5-3] 구매처별 거래 비중



- 데이터서비스 세부분야 중 데이터 거래 분야는 중소기업의 구매가 49.3%, 정부 및 공공분야 구매가 27.9%로 나타났고, 데이터 분석 제공 분야는 대부분 중소기업(60.0%)과 대기업(40.0%)에서 구매하는 경향을 보이고 있다.

[표5-3] 데이터서비스 분야별 구매처별 거래 비중

구분	응답수	대기업	중소기업	정부·공공	개인
데이터서비스 전체	(41)	27.6	48.3	15.8	8.3
데이터 거래	(14)	17.1	49.3	27.9	5.7
정보 제공	(23)	31.7	45.7	11.3	11.3
데이터 분석 제공	(4)	40.0	60.0	0.0	0.0

※ 데이터 구매가 있다고 응답한 기업에 한해 조사된 결과임

(단위: 개, %)

※ 무응답 제외

○ 데이터를 구매하는 경우, 적절한 가격으로 구매한다는 의견이 96.6%로 가장 높았다.

[표5-4] 구매 거래 가격 형태

구분	완전 무료	광고/ 홍보 조건 무료	파격 할인된 가격	적정한 가격	기타	무응답
응답수	(2)	(2)	(0)	(114)	(0)	(0)
비중	1.7	1.7	0.0	96.6	0.0	0.0

※ 데이터 구매가 있다고 응답한 기업에 한해 조사된 결과임

(단위: 개, %)

○ 구매처 유형별 구매 거래 가격 형태를 보면, 적정가격으로 구매한다는 의견이 가장 많았다. 중소기업에서 광고 및 홍보를 조건으로 무료로 제공받는(3.8%) 경우도 있었으며, 정부 및 공공 기관 데이터는 완전 무료로 구매한다는 의견도 11.1%로 나타났다.

[표5-5] 데이터서비스 분야별 구매처 유형별 거래 가격 형태

구분		응답수	완전 무료	광고/ 홍보 조건 무료	파격 할인된 가격	적정한 가격	기타	무응답
대기업	데이터서비스 전체	(17)	0.0	0.0	0.0	100.0	0.0	0.0
	데이터 거래	(5)	0.0	0.0	0.0	100.0	0.0	0.0
	정보 제공	(10)	0.0	0.0	0.0	100.0	0.0	0.0
	데이터 분석 제공	(2)	0.0	0.0	0.0	100.0	0.0	0.0
중소 기업	데이터서비스 전체	(26)	0.0	3.8	0.0	96.2	0.0	0.0
	데이터 거래	(9)	0.0	0.0	0.0	100.0	0.0	0.0
	정보 제공	(14)	0.0	7.1	0.0	92.9	0.0	0.0
	데이터 분석 제공	(3)	0.0	0.0	0.0	100.0	0.0	0.0
정부· 공공	데이터서비스 전체	(9)	11.1	0.0	0.0	88.9	0.0	0.0
	데이터 거래	(5)	20.0	0.0	0.0	80.0	0.0	0.0
	정보 제공	(4)	0.0	0.0	0.0	100.0	0.0	0.0
개인	데이터서비스 전체	(7)	0.0	0.0	0.0	100.0	0.0	0.0
	데이터 거래	(3)	0.0	0.0	0.0	100.0	0.0	0.0
	정보 제공	(4)	0.0	0.0	0.0	100.0	0.0	0.0

※ 데이터 구매가 있다고 응답한 기업에 한해 조사된 결과임

(단위: 개, %)

- 데이터 판매처별 거래 비중을 살펴보면 중소기업이 35.6%로 가장 높고, 대기업이 23.9%, 개인 23.4%, 정부 및 공공기관은 17.1%의 순으로 나타났다.

[그림5-4] 판매처별 거래 비중



- 데이터서비스 분야 중 데이터 거래와 정보제공 기업은 정부 및 공공기관을 판매처로 둔 비중이 16% 이상으로 나타났다.

[표5-6] 데이터서비스 분야별 판매처별 거래 비중

구분	응답수	대기업	중소기업	정부·공공	개인
데이터서비스 전체	(74)	23.9	35.6	17.1	23.4
데이터 거래	(23)	20.8	53.2	16.2	9.8
정보 제공	(49)	26.4	24.7	18.2	30.7
데이터 분석 제공	(2)	0.0	100.0	0.0	0.0

※ 데이터 판매가 있다고 응답한 기업에 한해 조사된 결과임

(단위: 개, %)

※ 무응답 제외

- 판매거래 가격 형태는 대부분 적절한 가격으로 거래가 이뤄지고 있다는 의견이 85.7%로 대부분을 차지하였다. 그 외에는 광고/홍보 조건 무료 4.3%, 파격 할인된 가격 3.7%, 완전 무료 2.5%로 나타났다.

[표5-7] 판매 거래 가격 형태

구분	완전 무료	광고/ 홍보 조건 무료	파격 할인된 가격	적정한 가격	기타	무응답
응답수	(8)	(14)	(12)	(276)	(2)	(10)
비중	2.5	4.3	3.7	85.7	0.6	3.2

※ 데이터 판매가 있다고 응답한 기업에 한해 조사된 결과임

(단위: 개, %)

- 판매처별 거래 가격 형태를 보면, 거래처 유형과 상관없이 적절한 가격으로 거래가 이뤄진다는 의견이 가장 높게 나타났다.

[표5-8] 데이터서비스 분야별 판매처 유형별 거래 가격 형태

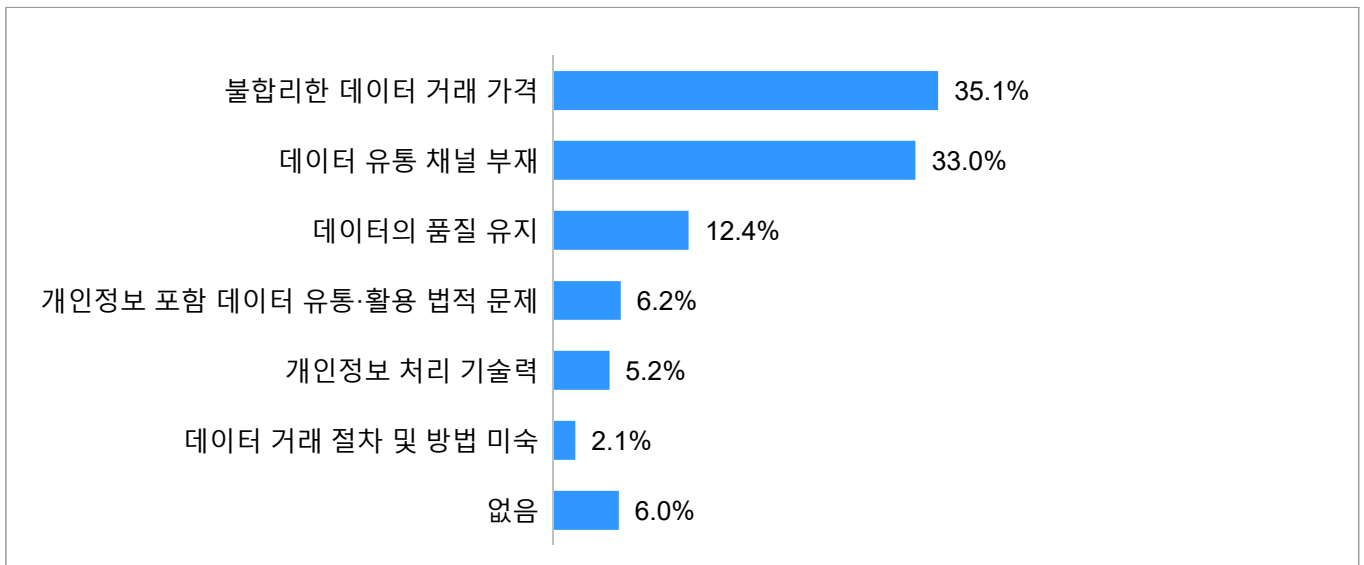
구분		응답수	완전 무료	광고/ 홍보 조건 무료	파격 할인된 가격	적정한 가격	기타	무응답
대기업	데이터서비스 전체	(42)	2.4	2.4	4.8	88.1	0.0	2.3
	데이터 거래	(15)	0.0	0.0	6.7	86.7	0.0	6.6
	정보 제공	(27)	3.7	3.7	3.7	88.9	0.0	0.0
중소 기업	데이터서비스 전체	(49)	2.0	6.1	2.0	85.7	2.0	2.2
	데이터 거래	(21)	0.0	0.0	0.0	95.2	0.0	4.8
	정보 제공	(26)	3.8	11.5	3.8	76.9	4.0	0.0
	데이터 분석 제공	(2)	0.0	0.0	0.0	100.0	0.0	0.0
정부· 공공	데이터서비스 전체	(32)	3.1	3.1	3.1	87.5	0.0	3.2
	데이터 거래	(12)	0.0	0.0	0.0	91.7	0.0	8.3
	정보 제공	(20)	5.0	5.0	5.0	85.0	0.0	0.0
개인	데이터서비스 전체	(38)	2.6	5.3	5.3	81.6	0.0	5.2
	데이터 거래	(11)	0.0	0.0	9.1	81.8	0.0	9.1
	정보 제공	(27)	3.7	7.4	3.7	81.5	0.0	3.7

※ 데이터 판매가 있다고 응답한 기업에 한해 조사된 결과임

(단위: 개, %)

- 데이터 거래 시 애로사항으로는 불합리한 데이터 거래 가격이 35.1%, 데이터 유통 채널 부재가 33.0%로 가장 많았다.
- 다음으로 데이터의 품질 유지가 어렵다가 12.4%, 개인정보 포함 데이터 유통·활용 법적 문제 6.2%, 개인정보 처리 기술력 5.2%, 데이터 거래 절차 및 방법 미숙 2.1% 순으로 나타났다.

[그림5-5] 데이터 거래 시 애로사항



- 데이터 분석 제공 분야는 데이터 유통 채널 부재와 개인정보 포함 데이터 유통·활용 법적 문제가 비교적 높게 나타났다.

[표5-9] 데이터서비스 분야별 데이터 거래 시 애로사항

구분	응답수	불합리한 데이터 거래 가격	데이터 유통 채널 부재	데이터의 품질 유지	개인정보 (일부)포함 데이터 유통· 활용 시 법적 문제	개인정보 처리 기술력	데이터 거래 절차 및 방법 미숙	없음
데이터서비스 전체	(41)	35.1	33.0	12.4	6.2	5.2	2.1	6.0
데이터 거래	(14)	48.5	21.2	12.1	9.1	3.0	0.0	6.1
정보 제공	(23)	29.6	37.0	14.8	1.9	5.6	3.7	7.4
데이터 분석 제공	(4)	20.0	50.0	0.0	20.0	10.0	0.0	0.0

※ 데이터 거래가 있다고 응답한 기업에 한해 조사된 결과임. 우선순위 가중치 적용

(단위: 개, %)

2) 데이터 유통·활용 활성화를 위한 필요사항

- 데이터서비스 기업들은 데이터 유통 및 활용 비즈니스 활성화를 위해 보유 데이터의 적절한 가치 평가(38.0%)를 가장 필요한 지원 사항이라고 응답했다. 다음으로 데이터 유통/홍보 채널(28.5%), 데이터 구입비 지원(13.1%) 등의 순으로 나타났다.

[표5-10] 데이터 유통·활용 활성화 필요지원 사항

구분	응답수	데이터 구입비 지원	보유 데이터의 적절한 가치 평가	데이터 유통/홍보 채널	데이터 유통 관련 법/기술 교육	없음
데이터서비스 전체	(382)	13.1	38.0	28.5	8.1	12.3
데이터 거래	(53)	18.9	35.8	32.1	5.7	7.5
정보 제공	(306)	13.1	37.6	28.1	8.2	13.0
데이터 분석 제공	(23)	0.0	47.8	26.1	13.0	13.1

(단위: 개, %)

- 데이터 거래 경험과 관계없이 보유 데이터의 적절한 가치 평가가 데이터 유통·활용 활성화를 위해 가장 필요한 지원 사항으로 나타났다.

[표5-11] 거래경험 유무별 데이터 유통·활용 활성화 필요지원 사항

구분	응답수	데이터 구입비 지원	보유 데이터의 적절한 가치 평가	데이터 유통/홍보 채널	데이터 유통 관련 법/기술 교육	없음
데이터서비스 전체	(382)	13.1	38.0	28.5	8.1	12.3
거래경험 있음	(110)	18.2	39.1	30.9	6.4	5.4
거래경험 없음	(272)	11.0	37.5	27.6	8.8	15.1

(단위: 개, %)

- 개인정보가 포함된 데이터를 유통하거나 활용하는 경우, 데이터 유통 및 활용 관련 법제도 개선이 42.9%로 가장 높고, 다음으로 개인정보 활용에 대한 인식 개선(22.0%), 개인정보 비식별화 등 처리기술 지원(14.9%), 등의 순으로 조사되었다.

[표5-12] 개인정보가 포함된 데이터 유통·활용 필요지원 사항

구분	응답수	개인정보 등 데이터 유통/ 활용 관련 법제도 개선	개인정보 비식별화 등 처리기술 지원	개인정보 활용에 대한 인식 개선	개인정보를 유통하는 별도의 플랫폼 개발	없음
데이터서비스 전체	(382)	42.9	14.9	22.0	5.2	15.0
데이터 거래	(53)	30.2	20.8	30.2	7.5	11.3
정보 제공	(306)	44.4	14.1	20.9	4.9	15.7
데이터 분석 제공	(23)	52.2	13.0	17.4	4.3	13.1

(단위: 개, %)

- 데이터 거래 경험 유무와 상관없이 개인정보 등 데이터 유통·활용 관련 법제도 개선을 가장 필요로 하였으며, 데이터 거래 경험이 있는 기업의 경우, 개인정보 비식별화 등 처리기술 지원(25.5%)이 데이터 거래 경험이 없는 기업에 비해 상대적으로 높게 나타났다.

[표5-13] 거래경험 유무별 개인정보가 포함된 데이터 유통·활용 필요지원 사항

구분	응답수	개인정보 등 데이터 유통/ 활용 관련 법제도 개선	개인정보 비식별화 등 처리기술 지원	개인정보 활용에 대한 인식 개선	개인정보를 유통하는 별도의 플랫폼 개발	없음
데이터서비스 전체	(382)	42.9	14.9	22.0	5.2	15.0
거래경험 있음	(110)	32.7	25.5	24.5	6.4	10.9
거래경험 없음	(272)	47.1	10.7	21.0	4.8	16.4

(단위: 개, %)

5.2 데이터 보유 현황

1) 보유 · 제공 DB 현황

○ 데이터서비스 기업의 DB 최초 개발 시기는 2005년 이후가 46.4%로 절반 가량을 차지하고 있다.

[표5-14] DB 최초 개발 시기

구분	DB수	1999년도 이전	2000년~2004년	2005년~2009년	2010년~2014년	2015년~2017년	모름
데이터서비스 기업	(330)	14.2	19.1	19.1	20.6	6.7	20.3

※ 데이터서비스 327개 기업이 응답한 결과임

(단위: 개, %)

○ 데이터를 활용하거나 외부 서비스를 하기 위해 구축한 DB의 평균 개발 기간은 13.8개월로 조사되었으며, 평균 DB 구축비용은 약 2.6억 원으로 나타났다.

[표5-15] DB 개발 완료 소요기간 및 구축 비용

구분	DB수	평균 DB 개발 소요기간	DB 구축비용	
			평균	합계
데이터서비스 기업	(330)	13.8	2.6	214

※ 데이터서비스 267개 기업이 응답한 결과임

(단위: 개, 개월, 억 원)

○ DB 개발 시 투입인력 규모는 DB 당 평균 8.3명이며, 전문 역량 수준별 투입인력을 보면, 중급 3.7명, 고급 3.6명으로 나타났다.

[표5-16] DB 개발 투입인원

구분	DB수	초급		중급		고급		특급		합계	
		평균	합계	평균	합계	평균	합계	평균	합계	평균	합계
데이터서비스 기업	(330)	0.8	92	3.7	446	3.6	435	0.2	30	8.3	1,003

※ 데이터서비스 121개 기업이 응답한 결과임

(단위: 개, 명)

- 구축된 DB 갱신 주기를 살펴보면, 실시간 갱신이 34.2%로 가장 많았으며, 매일(23.0%), 연간(14.2%), 매주(8.2%), 매월(5.5%) 순으로 나타났다.

[표5-17] 갱신 주기

구분	DB수	실시간	매일	매주	격주	매월	격월	분기별	연간	모름
데이터서비스 기업	(330)	34.2	23.0	8.2	2.1	5.5	0.3	3.6	14.2	8.9

※ 데이터서비스 327개 기업이 응답한 결과임

(단위: 개, %)

- DB 갱신 소요비용은 평균 약 1,100만 원 정도 소요되는 것으로 나타났다.

[표5-18] 갱신 소요비용

구분	DB수	100만원 미만	200만원 미만	500만원 미만	500만원 이상	평균 (억 원)
데이터서비스 기업	(330)	62.3	3.8	5.6	28.3	0.11

※ 데이터서비스 53개 기업이 응답한 결과임

(단위: 개, %, 억 원)

- DB 갱신을 위해 평균 6.5명의 인력이 투입되고 있다.

[표5-19] 갱신 투입인력

구분	DB수	초급		중급		고급		특급		합계	
		평균	합계	평균	합계	평균	합계	평균	합계	평균	합계
데이터서비스 기업	(330)	2.3	251	2.1	231	2.2	244	0.0	1	6.5	727

※ 데이터서비스 111개 기업이 응답한 결과임

(단위: 개, 명)

2) 보유·제공 DB 특징

- DB의 데이터 유형을 살펴보면, 비정형 데이터와 정형 데이터가 혼재된 비중이 41.5%로 가장 높은 것으로 조사되었다.

[표5-20] DB 수록데이터 유형

구분	DB수	비정형 데이터	정형 데이터	비정형+정형 혼재	모름
데이터서비스 기업	(330)	28.2	23.0	41.5	7.3

※ 데이터서비스 327개 기업이 응답한 결과임

(단위: 개, %)

- 데이터상품의 성격을 보면, 검색이 가능한 계층적 구조의 연관 데이터의 집합이라는 의견이 56.7%로 가장 많은 것으로 나타났다.

[표5-21] 데이터 상품 성격

구분	DB수	데이터 목록	검색이 가능한 계층적 구조의 연관 데이터 집합	이용자 행동 데이터 (모니터링)	산업분야 별 트래킹 데이터	데이터 탑재형 SW/HW	기타	모름
데이터서비스 기업	(330)	23.3	56.7	4.5	4.2	2.7	0.9	7.7

※ 데이터서비스 327개 기업이 응답한 결과임

(단위: 개, %)

- DB 유통 방식은 유료 정보서비스 제공이 42.1%로 가장 많고, 무료 정보서비스 제공은 40.6%로 나타났다.

[표5-22] DB 유통 방식

구분	DB수	무료 정보서비스 제공	유료 정보서비스 제공	B2B간 원천데이터 /DB 유료 판매	온라인 유통 플랫폼을 통한 유료 판매	내부 활용목적 비공개	기타	모름
데이터서비스 기업	(330)	40.6	42.1	2.7	1.8	6.7	0.3	5.8

※ 데이터서비스 327개 기업이 응답한 결과임

(단위: 개, %)

The background features a stylized city skyline with various skyscrapers in shades of blue and grey. Overlaid on this are large, light blue geometric shapes, including a prominent diagonal band and several triangular sections, creating a modern, architectural feel.

부록1. 조사표



통계법 제33조(비밀의 보호 등) 통계 작성과정에서 알려진 사항으로서 개인 또는 법인이나 단체의 비밀에 속하는 사항은 보호되어야 한다.

주관기관 : 과학기술정보통신부
전담기관 : 한국데이터진흥원

2017년 데이터산업현황조사 질의서

안녕하십니까?

데이터산업 현황 조사는 DB산업육성사업의 일환으로 데이터산업의 시장 규모 및 전문 인력 현황, 데이터 보유현황, 데이터 유통 현황 등 데이터산업 전반에 대한 현황을 수집해 데이터산업 기초 통계를 작성하고자 과학기술정보통신부와 한국데이터진흥원이 실시하는 조사입니다.

본 조사 결과는 DB산업 육성·지원 정책 수립뿐만 아니라 국가 정책 수립의 중요한 근거 자료로 활용되며, 이 외의 다른 목적으로는 사용하지 않습니다.

바쁘신 중에라도 잠시 시간을 내시어 응답해 주시기를 부탁드립니다.

아울러 조사 결과는 분석 보고서, 데이터산업 백서를 통하여 제공할 계획입니다.
성실한 응답을 부탁드립니다. 귀사의 건승과 일익 번창하심을 기원합니다.

본 조사에 협조해주셔서 감사합니다.

2017년 10월

한국데이터진흥원

© 수행기관 : (주)마크로밀엠브레인

[응답자 프로파일]

회사명		사업자등록번호	
회사대표번호		설립연도	
회사홈페이지	http://	기업유형	☐ 단독사업체 ☐ 본사 ☐ 지사
주소			

응답자 성명		부서/직급	
전화번호		핸드폰 번호	
이메일			

I. 데이터 관련 사업 분야와 매출 관련 질문

문1. 귀사에서 하고있는 데이터 산업 관련 사업 분야는 다음 중 무엇입니까?

해당되는 것을 **모두** 말씀해 주세요.

- ① 데이터 솔루션 ② 데이터 구축/컨설팅 ③ 데이터 서비스
④ 데이터 인프라 ⑤ 해당 없음 → 조사 종료

■ 데이터 산업 정의

구분	설명	제외영역
데이터 솔루션	- DBMS, DBMS관리, 데이터 모델링, 분석·시각화, 검색엔진, 품질 등 관련 솔루션 제품으로 비즈니스를 영위하는 기업 - 라이선스, 유지보수, 커스터마이징(개발)에서 매출 발생	한글, OS 등 오피스 및 컴퓨터 운영프로그램, 게임, 이미지 뷰어/렌더링 프로그램, 유틸리티, 영상/음향 재생 프로그램 등 제외
데이터 구축	DB설계, 데이터 이행 등을 포함한 DB시스템 구축, 문서·영상·영상 등의 데이터를 DB로 변환·정비하는 데이터 처리, 데이터 외부 제공을 위한 API, LOD 구축, DW, Data Lake 등의 데이터 구축 등의 비즈니스를 영위하는 기업(데이터/DB 관련 SI·IT아웃소싱 포함)	네트워크 시스템 구축, H/W 유통 및 전산실 구축, IT아웃소싱(유지보수, 인력파견, 시스템운영), 공장/업무자동화 시스템 구축, 홈페이지 구축 등 제외
데이터 컨설팅	데이터 거버넌스, 품질, 데이터 설계 및 데이터 활용 등 데이터 관련 기획 및 컨설팅 비즈니스를 영위하는 기업	교육, 일반 경영컨설팅 제외
데이터 서비스	데이터를 활용해 정보제공, 데이터 거래, 분석 결과 정보 등을 온·오프라인(모바일, 앱 등 포함)으로 제공하면서 데이터 이용료/수수료 또는 광고료 등으로 비즈니스를 영위하거나, 마케팅을 목적으로 데이터를 수집·가공하여 판매하는 기업(데이터 브로커)	정보전달을 주목적으로 하며, 온라인게임 등 오락적 목적이거나 전자상거래의 일반재화 판매매출 등은 제외
데이터 인프라	데이터의 저장과 활용을 목적으로 하는 서버, 스토리지, 네트워크 등 기반 구조를 제조, 판매하거나 기획, 구축, 운영 또는 관련 컨설팅 비즈니스를 영위하는 기업	DBMS가 탑재되지 않는 서버와 데이터 인프라 관련 시설물 등은 제외

문2. [문1에서 선택한 데이터 산업 분야(대분류)에서만 응답]

귀사의 데이터 관련 사업 분야(주사업 및 부사업)에 대해 아래 표와 부록의 분야별 정의를 참고하여 **모두 체크**해주시시오.(중복체크) ※ [부록1] 2017년 데이터산업분류 상세 정의 참고

주사업(택1) (전체 매출 비중 1위)	부사업 (중복 선택)
--------------------------	----------------

대분류	중분류	소분류			
데이터 솔루션	데이터 수집	☐	① 로그데이터 수집	☐	② 웹데이터 수집
		☐	③ 데이터 검색		
	데이터 설계	☐	④ 데이터 아키텍처	☐	⑤ 데이터 모델링
		☐	⑥ 디스크 DBMS	☐	⑦ 인메모리 DBMS
	DBMS	☐	⑧ DBMS 어플라이언스	☐	⑨ 기타 DBMS
		☐	⑩ DB 운영 관리	☐	⑪ DB 성능 관리
	데이터 관리	☐	⑫ DB 보안 관리	☐	⑬ 데이터 분산 처리
		☐	⑭ 데이터 통합/흐름관리		
	데이터 품질	☐	⑮ 데이터 품질 관리	☐	⑯ 메타데이터 관리
		☐	⑰ 마스터데이터 관리		
	데이터 분석	☐	⑱ 정형 데이터 분석	☐	⑲ 비정형 데이터 분석
		☐	⑳ 실시간 데이터 분석	☐	㉑ 데이터 시각화 분석
	데이터 플랫폼	☐	㉒ 데이터 플랫폼		

대분류	중분류	소분류			
데이터 구축/컨설팅	구축	☐	㉓ DB설계·구축	☐	㉔ 데이터 이행·처리
		☐	㉕ 기계처리형 데이터 구축		
	컨설팅	☐	㉖ 데이터 설계	☐	㉗ 데이터 품질
		☐	㉘ 데이터 관리 성능개선	☐	㉙ 데이터 거버넌스 체계
		☐	㉚ 데이터 분석		
데이터 서비스	데이터 거래	☐	㉛ 데이터 중개	☐	㉜ 데이터 판매
		☐	㉝ 데이터 신디케이션		
	정보 제공	☐	㉞ 교육/훈련	☐	㉟ 신용/재무
		☐	㊱ 통계	☐	㊲ 경영/비즈니스
		☐	㊳ 취업/창업	☐	㊴ 뉴스
		☐	㊵ 포털	☐	㊶ 행정/법률
		☐	㊷ 학술	☐	㊸ 문화/예술 ☐ ㊹ 생활
	데이터 분석 제공	☐	㊺ 환경분석서비스	☐	㊻ 예측분석서비스
		☐	㊼ 심사분석서비스	☐	㊽ 기타 데이터분석서비스

■ [문1 ③ 데이터 서비스 해당 기업일 경우 참조] 데이터 서비스 - 데이터 판매와 정보제공 구분

※ 최종이용자에게 정보전달을 목적으로 데이터를 가공해 무료 또는 서비스이용료를 부과해 제공하는 서비스는 “정보서비스”로 선택해주시요. B2B로 원천데이터를 판매하는 경우에는 “데이터거래/데이터판매”로 선택해주시요.
(예: 학술논문검색 및 원문다운로드 = 학술정보서비스 / 학술논문 목록데이터 판매 = 데이터판매)

문3. 귀사의 데이터 산업 분야의 **매출액**을 기입해주시요.

구분	2016년	2017년 6월(상반기)	2017년(예상)
데이터 솔루션 분야 매출	백만원	백만원	백만원
데이터 구축/컨설팅 분야 매출	백만원	백만원	백만원
데이터 서비스 분야 매출	백만원	백만원	백만원
합계	백만원	백만원	백만원

■ [문1 ③ 데이터 서비스 해당 기업일 경우 참조] 데이터 서비스 수익기반별 매출 유형

※ 데이터 서비스 전체 매출은 아래 설명과 같이 데이터 서비스 매출과 광고 매출을 포함한 금액입니다.

구분	기준
데이터서비스	DB로 구축된 대량 데이터를 기업이나 기관 등에 판매하면서 창출되는 매출(데이터 판매)
	특화된 데이터분야를 DB로 구축한 뒤 개인 또는 기업을 대상으로 검색 및 조회서비스를 제공하며 일정 수수료가 발생하는 매출(신용평가정보 이용료 등)
	기존 데이터를 수요자의 요구에 맞춰 가공·편집하여 제공할 때 발생하는 수수료
	데이터 등록 수수료(구인구직정보 등록료, 부동산 매물정보 등록료)
	데이터 중개, 포털 등에서 데이터 제공자로부터 데이터를 제공받아 판매한 데이터 수익(중개 수수료, 플랫폼 이용료 등)
	SW판매 또는 SI사업 수행시 추가되는 데이터로서 DB 판매가 이뤄지는 경우의 데이터 판매 매출(네비게이션 제조사의 지리데이터 및 위치데이터 판매 등)
광고	포털사 등의 인터넷 광고를 통한 광고 수익, 데이터서비스 외 부가적으로 발생하는 광고 수익

문4. [문2에서 선택한 데이터 산업(소분류)에 대한 것만 응답]

귀사의 데이터 관련 사업 매출 실적을 기입해주시오.

문4-1. 데이터 솔루션 분야인 경우 (2016년, 2017년)

데이터 솔루션		2016년 실적	각 영역별 매출 비중				업종별 매출 비중						
중분류	소분류		라이선스	유지보수	개발	합계	공공	금융	제조/건설	유통/서비스	통신/미디어	의료	합계
데이터 수집	데이터 검색	백만원	%	%	%	100%	%	%	%	%	%	%	100%
	로그데이터 수집	백만원	%	%	%	100%	%	%	%	%	%	%	100%
	웹데이터 수집	백만원	%	%	%	100%	%	%	%	%	%	%	100%
데이터 설계	데이터 아키텍처	백만원	%	%	%	100%	%	%	%	%	%	%	100%
	데이터 모델링	백만원	%	%	%	100%	%	%	%	%	%	%	100%
DBMS	디스크 DBMS	백만원	%	%	%	100%	%	%	%	%	%	%	100%
	인메모리 DBMS	백만원	%	%	%	100%	%	%	%	%	%	%	100%
	DBMS 어플라이언스	백만원	%	%	%	100%	%	%	%	%	%	%	100%
	기타 DBMS	백만원	%	%	%	100%	%	%	%	%	%	%	100%
데이터 관리	DB 운영 관리	백만원	%	%	%	100%	%	%	%	%	%	%	100%
	DB 성능 관리	백만원	%	%	%	100%	%	%	%	%	%	%	100%
	DB 보안 관리	백만원	%	%	%	100%	%	%	%	%	%	%	100%
	데이터 분산 처리	백만원	%	%	%	100%	%	%	%	%	%	%	100%
	데이터 통합/흐름 관리	백만원	%	%	%	100%	%	%	%	%	%	%	100%
데이터 품질	데이터 품질 관리	백만원	%	%	%	100%	%	%	%	%	%	%	100%
	메타데이터 관리	백만원	%	%	%	100%	%	%	%	%	%	%	100%
	마스터데이터 관리	백만원	%	%	%	100%	%	%	%	%	%	%	100%
데이터 분석	정형데이터 분석	백만원	%	%	%	100%	%	%	%	%	%	%	100%
	비정형데이터 분석	백만원	%	%	%	100%	%	%	%	%	%	%	100%
	실시간데이터 분석	백만원	%	%	%	100%	%	%	%	%	%	%	100%
	데이터 시각화 분석	백만원	%	%	%	100%	%	%	%	%	%	%	100%
데이터 플랫폼	데이터 플랫폼	백만원	%	%	%	100%	%	%	%	%	%	%	100%
합계		백만원	-			100%	-						100%

* 합계는 문3의 2016년 데이터 솔루션 전체 매출액과 동일해야 함

데이터 솔루션		2017년 6월 실적 (상반기)	2017년 예상 실적 (전체)	각 영역별 매출 비중				업종별 매출 비중							
중분류	소분류			라이 선스	유지 보수	개발	합계	공공	금융	제조/ 건설	유통/ 서비스	통신/ 미디어	의료	합계	
데이터 수집	데이터 검색	백만원	백만원	%	%	%	100%	%	%	%	%	%	%	100%	
	로그데이터 수집	백만원	백만원	%	%	%	100%	%	%	%	%	%	%	100%	
	웹데이터 수집	백만원	백만원	%	%	%	100%	%	%	%	%	%	%	100%	
데이터 설계	데이터 아키텍처	백만원	백만원	%	%	%	100%	%	%	%	%	%	%	100%	
	데이터 모델링	백만원	백만원	%	%	%	100%	%	%	%	%	%	%	100%	
DBMS	디스크 DBMS	백만원	백만원	%	%	%	100%	%	%	%	%	%	%	100%	
	인메모리 DBMS	백만원	백만원	%	%	%	100%	%	%	%	%	%	%	100%	
	DBMS 어플라이언스	백만원	백만원	%	%	%	100%	%	%	%	%	%	%	100%	
	기타 DBMS	백만원	백만원	%	%	%	100%	%	%	%	%	%	%	100%	
데이터 관리	DB 운영 관리	백만원	백만원	%	%	%	100%	%	%	%	%	%	%	100%	
	DB 성능 관리	백만원	백만원	%	%	%	100%	%	%	%	%	%	%	100%	
	DB 보안 관리	백만원	백만원	%	%	%	100%	%	%	%	%	%	%	100%	
	데이터 분산 처리	백만원	백만원	%	%	%	100%	%	%	%	%	%	%	100%	
	데이터 통합/흐름관리	백만원	백만원	%	%	%	100%	%	%	%	%	%	%	100%	
데이터 품질	데이터 품질 관리	백만원	백만원	%	%	%	100%	%	%	%	%	%	%	100%	
	메타데이터 관리	백만원	백만원	%	%	%	100%	%	%	%	%	%	%	100%	
	마스터데이터 관리	백만원	백만원	%	%	%	100%	%	%	%	%	%	%	100%	
데이터 분석	정형데이터 분석	백만원	백만원	%	%	%	100%	%	%	%	%	%	%	100%	
	비정형데이터 분석	백만원	백만원	%	%	%	100%	%	%	%	%	%	%	100%	
	실시간데이터 분석	백만원	백만원	%	%	%	100%	%	%	%	%	%	%	100%	
	데이터 시각화 분석	백만원	백만원	%	%	%	100%	%	%	%	%	%	%	100%	
데이터 플랫폼	데이터 플랫폼	백만원	백만원	%	%	%	100%	%	%	%	%	%	%	100%	
합계		백만원	백만원	-			100%	-							100%

* 합계는 문3의 2017년 데이터 솔루션 전체 예상 매출액과 동일해야 함

[문2에서 선택한 데이터 산업(소분류)에 대한 것만 응답]

문4-2. 데이터 구축/컨설팅 분야인 경우 (2016년, 2017년)

데이터 구축/컨설팅		2016년 실적	업종별 매출 비중						
중분류	소분류		공공	금융	제조/건설	유통/서비스	통신/미디어	의료	합계
구축	DB설계·구축	백만원	%	%	%	%	%	%	100%
	데이터 이행·처리	백만원	%	%	%	%	%	%	100%
	기계처리형 데이터 구축	백만원	%	%	%	%	%	%	100%
컨설팅	데이터 설계	백만원	%	%	%	%	%	%	100%
	데이터 품질	백만원	%	%	%	%	%	%	100%
	데이터 관리 성능개선	백만원	%	%	%	%	%	%	100%
	데이터 거버넌스 체계	백만원	%	%	%	%	%	%	100%
	데이터 분석	백만원	%	%	%	%	%	%	100%
합계		백만원	-						100%

* 합계는 문3의 2016년 데이터 구축/컨설팅 전체 매출액과 동일해야 함

데이터 구축/컨설팅		2017년 6월 실적 (상반기)	2017년 예상 실적 (전체)	업종별 매출 비중						
중분류	소분류			공공	금융	제조/건설	유통/서비스	통신/미디어	의료	합계
구축	DB설계·구축	백만원	백만원	%	%	%	%	%	%	100%
	데이터 이행·처리	백만원	백만원	%	%	%	%	%	%	100%
	기계처리형 데이터 구축	백만원	백만원	%	%	%	%	%	%	100%
컨설팅	데이터 설계	백만원	백만원	%	%	%	%	%	%	100%
	데이터 품질	백만원	백만원	%	%	%	%	%	%	100%
	데이터 관리 성능개선	백만원	백만원	%	%	%	%	%	%	100%
	데이터 거버넌스 체계	백만원	백만원	%	%	%	%	%	%	100%
	데이터 분석	백만원	백만원	%	%	%	%	%	%	100%
합계		백만원	백만원	-						100%

* 합계는 문3의 2017년 데이터 구축/컨설팅 전체 예상 매출액과 동일해야 함

[문2에서 선택한 데이터 산업(소분류)에 대한 것만 응답]

문4-3. 데이터 서비스 분야인 경우 (2016년, 2017년)

구분		2016년 실적	서비스 기반별 매출 비중				수익 기반별 매출 비중		
			PC	모바일	오프라인	합계	데이터 서비스	광고	합계
데이터 거래	데이터 중개	백만원	%	%	%	100%	%	%	100%
	데이터 판매	백만원	%	%	%	100%	%	%	100%
	데이터 신디케이션	백만원	%	%	%	100%	%	%	100%
정보 제공	교육/훈련	백만원	%	%	%	100%	%	%	100%
	신용/재무	백만원	%	%	%	100%	%	%	100%
	통계	백만원	%	%	%	100%	%	%	100%
	경영/비즈니스	백만원	%	%	%	100%	%	%	100%
	취업/창업	백만원	%	%	%	100%	%	%	100%
	뉴스	백만원	%	%	%	100%	%	%	100%
	포털	백만원	%	%	%	100%	%	%	100%
	행정/법률	백만원	%	%	%	100%	%	%	100%
	학술	백만원	%	%	%	100%	%	%	100%
	문화/예술	백만원	%	%	%	100%	%	%	100%
	생활	백만원	%	%	%	100%	%	%	100%
데이터 분석 제공	환경분석서비스	백만원	%	%	%	100%	%	%	100%
	예측분석서비스	백만원	%	%	%	100%	%	%	100%
	심사분석서비스	백만원	%	%	%	100%	%	%	100%
	기타 데이터분석서비스	백만원	%	%	%	100%	%	%	100%
기타 매출(데이터서비스 무관)		백만원	-						
합계*		백만원	-			100%	-		100%

* 합계는 문3의 2016년 데이터 서비스 전체 매출액과 동일해야 함

구분		2017년 6월 실적 (상반기)	2017년 예상 실적 (전체)	서비스 기반별 매출 비중				수익 기반별 매출 비중		
				PC	모바일	오프라인	합계	데이터 서비스	광고	합계
데이터 거래	데이터 중개	백만원	백만원	%	%	%	100%	%	%	100%
	데이터 판매	백만원	백만원	%	%	%	100%	%	%	100%
	데이터 신디케이션	백만원	백만원	%	%	%	100%	%	%	100%
정보 제공	교육/훈련	백만원	백만원	%	%	%	100%	%	%	100%
	신용/재무	백만원	백만원	%	%	%	100%	%	%	100%
	통계	백만원	백만원	%	%	%	100%	%	%	100%
	경영/비즈니스	백만원	백만원	%	%	%	100%	%	%	100%
	취업/창업	백만원	백만원	%	%	%	100%	%	%	100%
	뉴스	백만원	백만원	%	%	%	100%	%	%	100%
	포털	백만원	백만원	%	%	%	100%	%	%	100%
	행정/법률	백만원	백만원	%	%	%	100%	%	%	100%
	학술	백만원	백만원	%	%	%	100%	%	%	100%
	문화/예술	백만원	백만원	%	%	%	100%	%	%	100%
	생활	백만원	백만원	%	%	%	100%	%	%	100%
데이터 분석 제공	환경분석서비스	백만원	백만원	%	%	%	100%	%	%	100%
	예측분석서비스	백만원	백만원	%	%	%	100%	%	%	100%
	심사분석서비스	백만원	백만원	%	%	%	100%	%	%	100%
	기타 데이터분석서비스	백만원	백만원	%	%	%	100%	%	%	100%
기타 매출(데이터서비스 무관)		백만원	백만원	-						
합계*		백만원	백만원	-			100%	-		100%

* 합계는 문3의 2017년 데이터 서비스 전체 예상 매출액과 동일해야 함

II. 데이터직무 인력 현황에 관련 질문

문5. 귀사의 **전체 종사자 수**에 대해 기입해주시요.

전체 종사자수	남자	여자
명	명	명

문6. 귀사에서 종사하고 있는 **데이터직무 관련 현재 인력수**와 **추가로 필요한 인력수**에 대해 기입해주시요.

※ 1명이 여러 분야의 업무를 담당할 경우, 가장 주된 업무를 수행하는 분야로 기입하여 주십시오.

※ 필요인력은 채용계획과 무관하게 현 시점에서 귀사에서 필요한 인력수를 기입해 주시면 됩니다.

구분	전체 현재인력	성별 현재인력		기술등급별 현재인력			향후 3년('20) 데이터직무 필요인력	기술등급별 필요인력		
		남자	여자	초급	중급	고급		초급	중급	고급
① DA	명	명	명	명	명	명	명	명	명	명
② 데이터 개발자	명	명	명	명	명	명	명	명	명	명
③ 데이터 엔지니어	명	명	명	명	명	명	명	명	명	명
④ 데이터 분석가	명	명	명	명	명	명	명	명	명	명
⑤ DBA	명	명	명	명	명	명	명	명	명	명
⑥ 데이터 사이언티스트	명	명	명	명	명	명	명	명	명	명
⑦ 데이터 컨설턴트	명	명	명	명	명	명	명	명	명	명
⑧ 데이터 기획·마케터	명	명	명	명	명	명	명	명	명	명

■ 데이터 직무 구분

직무명	설명
①DA (data architect)	데이터를 기반으로 IT 정책, 표준화, 구조, 설계 및 이행을 하는 직무 개념적, 논리적, 물리적 데이터 설계 수행
②데이터 개발자	조직업무 기반 IT시스템 구축에서 DB 및 데이터를 이용하여 프로그래밍 하는 직무 Hadoop, NoSQL, MapReduce 등의 기술을 활용해 빅데이터 처리, LOD 구축 등 데이터 관련 프로그램 개발 포함
③데이터 엔지니어	DBMS, Hadoop, NoSQL 등 데이터 관련 제품에 대한 기술지원자, 제품 개발자, 유지보수 등의 직무
④데이터 분석가	정형, 비정형 데이터 등 다양한 데이터를 식별, 관리, 조작, 분석하여 기업 경영의 의사결정에 필요한 자료를 만들어내는 직무 통계, 머신러닝, 텍스트마이닝 기반 데이터 분석, 분석결과 시각화 등 포함
⑤DBA (database administrator)	DB관리체계와 자료를 검토, 개선하고 DB의 구성, 변경, 용량, 성능, 가용성, 보안, 장애, 문제관리 등 운영시스템의 관리 직무
⑥데이터 사이언티스트	조직 내외부 데이터의 관리/활용/분석 체계를 만들고, 분석을 통해 프로세스 혁신 및 신제품 개발, 마케팅 전략 결정 등의 의사결정을 이끌어내는 직무
⑦데이터 컨설턴트	성능튜닝, 데이터아키텍처, 문제해결 등을 총칭하는 컨설팅사의 직무 빅데이터 분석을 토대로 기업이 앞으로 나아갈 방향, 해결책 등을 제시하는 업무 포함
⑧데이터 기획·마케터	(빅)데이터 관련 제품/서비스 기획, 기술영업 및 마케팅 직무

■ 데이터 기술 등급 구분

※ 기술 등급별로 자격기준, 경력기준, 학력기준 요건 중 어느 한 기준만 충족되면 해당됨

기술등급	기준
초급	정보처리기사, 정보처리산업기사 등 동급 Data 관련 자격 취득자
	전문학사 및 학사학위 보유자
	고등학교 졸업 후 3년 이상 경력
중급	기사자격 또는 Data 관련 자격 취득 후 경력 3년 이상 (Data 관련 자격 : 데이터아키텍처전문가, SQL개발자, OCP, 데이터분석전문가, MCDBA, CSP, MySQL, DB Administrator, DB2 for Linux and Windows)
	산업기사 자격 취득 후 7년 이상 경력 석사의 경우 기사자격 또는 Data 관련 자격 취득 후 경력 2년 이상
	정보처리기사자격 또는 Data관련 자격 취득 후 3년이상 경력
고급	정보관리 기술사, 컴퓨터시스템응용 기술사
	중급 이후 3년 이상 경력
	Data 고급 자격 취득자 (Data고급자격 : 데이터아키텍처전문가, SQL전문가, OCM, 데이터 분석 전문가)
	박사의 경우 기사 및 Data관련 자격 취득자

문6-1. [빅데이터 매출이 있는 경우 & 문6 ②③④⑦⑧번 직무 응답자 대상]

귀사에서 다음의 직무에 종사하고 있는 인력 중 **빅데이터 관련 기술을 보유하고 있는 현재 인력 수와 추가로 필요한 인력 수(향후 3년)**에 대해 기입해주시시오.

※ 빅데이터 인력수는 5번 문항에서 입력한 현재인력 및 필요인력 총합계에 포함되어야 합니다.

(예: 데이터 엔지니어 현재인력 10명 중 5명이 빅데이터 관련 기술을 보유한 경우 "5명"으로 기입)

구분	빅데이터 관련 기술보유 현재인력 전체	성별 현재인력		기술등급별 현재인력			향후 3년('20) 빅데이터 관련 기술보유 필요인력	기술등급별 필요인력		
		남자	여자	초급	중급	고급		초급	중급	고급
② 빅데이터 개발자	명	명	명	명	명	명	명	명	명	명
③ 빅데이터 엔지니어	명	명	명	명	명	명	명	명	명	명
④ 빅데이터 분석가	명	명	명	명	명	명	명	명	명	명
⑦ 빅데이터 컨설턴트	명	명	명	명	명	명	명	명	명	명
⑧ 빅데이터 기획·마케터	명	명	명	명	명	명	명	명	명	명

■ 빅데이터 관련 기술 예

- Hadoop, NoSQL, MapReduce, Spark 등 빅데이터를 처리하거나 프로그램을 개발하기 위한 관련 기술 보유
- 두 개 이상의 데이터 포맷(예: 정형, 비정형)과 데이터 소스(예: 내부, 외부)를 활용하는 빅데이터 분석 기술 보유
- 실시간으로 스트리밍 되는 데이터 분석 기술 보유
- 빅데이터 관련 제품 및 서비스 기획 능력 보유
- 빅데이터 분석을 통한 문제해결, 의사결정 도출 및 제시 능력
- 기타 빅데이터 분석, 처리, 개발 등을 위한 관련 기술 보유 등

문7. 귀사에서 **데이터 전문인력 채용 시 우대하는 자격증**을 모두 체크해주시요.

※ 그 외 국내 자격증 및 해외 자격증은 “⑨기타”에 입력해주시요.

- ① 데이터아키텍처전문가(DAP)

② 데이터아키텍처 준전문가(DAsP)

③ SQL전문가(SQLP)

④ SQL개발자(SQLD)

⑤ 데이터 분석 전문가(ADP)

⑥ 데이터 분석 준전문가(ADsP)

⑦ 정보처리기사

⑧ 사회조사분석사1급

⑨ 사회조사분석사2급

⑩ 기타 : _____

▣ 국내 데이터 전문인력 자격증

자격증	내용
데이터아키텍처 전문가 (DAP)	효과적인 데이터아키텍처 구축을 위해 전사아키텍처와 데이터품질관리에 대한 지식을 바탕으로 데이터 요건분석, 데이터 표준화, 데이터 모델링, 데이터베이스 설계와 이용 등의 직무를 수행하는 전문가자격검정
데이터아키텍처 준전문가 (DAsP)	효과적인 데이터아키텍처 구축을 위해 전사 아키텍처에 대한 지식을 바탕으로 데이터 요건분석, 데이터 표준화, 데이터 모델링 등의 직무를 수행하는 실무자자격검정
SQL 전문가 (SQLP)	데이터를 조작하고 추출하는데 있어서 정확하고 최적의 성능을 발휘하는 SQL을 작성할 수 있고, 이를 토대로 SQL을 내포하는 데이터베이스 프로그램이나 응용 소프트웨어의 성능을 최적화하거나, 이러한 성능 최적화를 지원 할 수 있는 데이터베이스 개체(뷰, 인덱스 등)의 설계와 구현 등의 직무를 수행하는 전문가자격검정
SQL 개발자 (SQLD)	데이터베이스와 데이터 모델링에 대한 지식을 바탕으로 응용 소프트웨어를 개발하면서 데이터를 조작하고 추출하는데 있어서 정확하고 최적의 성능을 발휘하는 SQL을 작성할 수 있는 개발자자격검정
데이터 분석 전문가 (ADP)	데이터 이해 및 처리 기술에 대한 기본지식을 바탕으로 데이터 분석 기획, 데이터 분석, 데이터 시각화 업무를 수행하고 이를 통해 프로세스 혁신 및 마케팅 전략 결정 등의 과학적 의사결정을 지원할 수 있는 전문가자격검정
데이터 분석 준전문가 (ADsP)	데이터 이해에 대한 기본지식을 바탕으로 데이터 분석 기획 및 데이터 분석 등의 직무를 수행할 수 있는 실무자자격검정
정보처리기사	정보시스템의 생명주기 전반에 걸친 프로젝트 업무를 수행하는 직무로서 계획수립, 분석, 설계, 구현, 시험, 운영, 유지보수 등의 업무를 수행할 수 있는 전문가자격검정
사회조사분석사	다양한 사회정보의 수집·분석·활용을 담당하는 새로운 직종으로 기 업, 정당, 지방자치단체, 중앙정부 등 각종 단체의 시장조사 및 여론조사 등에 대한 계획을 수립하고 조사를 수행하며 그 결과를 분석, 보고서를 작성하는 전문가 자격검정

문8. 귀사에서 우대하는 **데이터 전문인력 자격증별**로 **채용 시 우대사항**을 모두 체크해주시요.

※ 빈칸에 우대사항 ①입사우대 ②응시료 지원 ③수당 지급 ④인사고과 반영 ⑤ 기타()를 입력해주시요.

구분		우대사항
(1) 자격증별 우대사항이 동일한 경우		① 입사 우대 ② 응시료 지원 ③ 수당 지급 ④ 인사고과 반영 ⑤ 기타()
(2) 자격증별 우대사항이 다른 경우	① 데이터아키텍처전문가(DAP)	
	② 데이터아키텍처 준전문가(DAsP)	
	③ SQL전문가(SQLP)	
	④ SQL개발자(SQLD)	
	⑤ 데이터 분석 전문가(ADP)	
	⑥ 데이터 분석 준전문가(ADsP)	
	⑦ 정보처리기사	
	⑧ 사회조사분석사1급	
	⑨ 사회조사분석사2급	
	⑩ 기타	

※ Part III은 문1 ③ 데이터 서비스 업체만 응답해 주십시오.

III. 데이터 보유 현황 관련 질문

문9. 귀사에서 서비스를 위해 보유하고 있거나, 제공 중인 DB 및 빅데이터에 대해 말씀해 주십시오.

주된 또는 핵심 DB 최대 3개까지 작성해주십시오.

※ 기업 내·외부의 모든 데이터를 일관된 체계로 구축, 운영할 수 있도록 제작한 모든 DB를 의미하며, 사업 추진 과정에서 자연스럽게 축적되어 분석·활용 가치가 있는 DB/빅데이터를 포함

※ 기존에 서비스 중이었으나 현재 서비스는 하지 않고 있지만 보유하고 있는 DB 또는 데이터도 포함

구분		주요 DB(1)				주요 DB(1)				주요 DB(3)			
DB명		예)분석용 이용자로그DB				예)뉴스DB, 논문DB,구인구직DB				예)특허정보, 법률정보DB			
해당DB 특허보유 현황 (특허 보유시 작성)		특허번호(출원/공개/등록) 특허명				특허번호(출원/공개/등록) 특허명				특허번호(출원/공개/등록) 특허명			
DB 설명 (수록 데이터에 대한 간략한 설명)													
데이터서비스(DB시스템)의 서비스 개시일		년		월		년		월		년		월	
데이터 구축부터 서비스 시작까지 구축 소요기간		년		개월		년		개월		년		개월	
데이터 구축부터 서비스 시작까지 구축 비용		백만원				백만원				백만원			
데이터 구축부터 서비스 시작까지 투입인력 수		초급	중급	고급	특급	초급	중급	고급	특급	초급	중급	고급	특급
		명	명	명	명	명	명	명	명	명	명	명	명
데이터가 차지하는 데이터 총량 또는 레코드 수		TB		건		TB		건		TB		건	
갱신주기 (해당 사항 체크(V)표시)		① 실시간 ③ 매주 ⑤ 매월 ⑦ 분기별	② 매일 ④ 격주 ⑥ 격월 ⑧ 연간			① 실시간 ③ 매주 ⑤ 매월 ⑦ 분기별	② 매일 ④ 격주 ⑥ 격월 ⑧ 연간			① 실시간 ③ 매주 ⑤ 매월 ⑦ 분기별	② 매일 ④ 격주 ⑥ 격월 ⑧ 연간		
갱신시 소요비용		백만원				백만원				백만원			
갱신시 투입인력 수 (초급/중급/고급/특급별 투입인원)		초급	중급	고급	특급	초급	중급	고급	특급	초급	중급	고급	특급
		명	명	명	명	명	명	명	명	명	명	명	명
DB 특징 (해당 사항 체크(V)표시)	수록데이터 유형	① 비정형 데이터 ② 정형 데이터 ③ 비정형 + 정형 혼재 ④ 기타 ()				① 비정형 데이터 ② 정형 데이터 ③ 비정형 + 정형 혼재 ④ 기타 ()				① 비정형 데이터 ② 정형 데이터 ③ 비정형 + 정형 혼재 ④ 기타 ()			
	데이터상품 성격	① 데이터 목록(contact list) ② 검색이 가능한 계층적 구조의 연관데이터 집합(directory) ③ 이용자 행동 데이터(모니터링) ④ 산업분야별 트래킹 데이터(예: 금융거래, 카드이용) ⑤ 데이터 탑재형 SW/HW ⑥ 기타()				① 데이터 목록(contact list) ② 검색이 가능한 계층적 구조의 연관데이터 집합(directory) ③ 이용자 행동 데이터(모니터링) ④ 산업분야별 트래킹 데이터(예: 금융거래, 카드이용) ⑤ 데이터 탑재형 SW/HW ⑥ 기타()				① 데이터 목록(contact list) ② 검색이 가능한 계층적 구조의 연관데이터 집합(directory) ③ 이용자 행동 데이터(모니터링) ④ 산업분야별 트래킹 데이터(예: 금융거래, 카드이용) ⑤ 데이터 탑재형 SW/HW ⑥ 기타()			
	유통 방식	① 무료 정보서비스 제공 ② 유료 정보서비스 제공 ③ B2B간 원천데이터/DB 유료 판매 ④ 온라인 유통 플랫폼을 통한 유료 판매 ⑤ 내부 활용목적 비공개 ⑥ 기타 ()				① 무료 정보서비스 제공 ② 유료 정보서비스 제공 ③ B2B간 원천데이터/DB 유료 판매 ④ 온라인 유통 플랫폼을 통한 유료 판매 ⑤ 내부 활용목적 비공개 ⑥ 기타 ()				① 무료 정보서비스 제공 ② 유료 정보서비스 제공 ③ B2B간 원천데이터/DB 유료 판매 ④ 온라인 유통 플랫폼을 통한 유료 판매 ⑤ 내부 활용목적 비공개 ⑥ 기타 ()			

※ Part IV는 문1 ③ 데이터 서비스 업체만 응답해 주십시오.

IV. 데이터 유통 관련 질문

문10. 귀사는 데이터를 거래한 경험이 있습니까?

- ① 데이터 판매 ② 데이터 구매 ③ 데이터 판매 및 구매 ④ 경험 없음

▣ 데이터 거래란?

개인이나 사업자가 신규 서비스 개발의 원천소재로 이용할 수 있도록 상품화하여 대량으로 거래/제휴하는 것을 말하며, 정보를 소비하는 수요자를 대상으로 인터넷 등을 통해 제공하는 정보서비스는 제외
예1) 티켓과 공연작품/공연정보DB를 결합한 서비스를 위한 데이터 거래 및 제휴
예2) 전문정보DB를 포털과 같은 대형 사이트에서 제공하기 위해 데이터 및 정보를 제공
예3) 마케팅에 활용하기 위해 기업 및 소비자 정보 DB 제공

문11. [거래 무경험자 대상] 귀사가 데이터 거래 경험이 없는 이유는 무엇입니까?

- ① 데이터 유통 채널 부재 ② 내부 인력 부족 ③ 필요한 데이터 자체 구축
④ 쓸만한 양질의 데이터 부재 ⑤ 사업의 특성상 거래 불필요 ⑥ 기타()

문12. [거래 경험자 대상] 귀사가 데이터를 거래할 때, 구매 또는 판매처별로 거래 현황을 기입해주시십시오.

※ '거래가격'란에는 아래의 '▣거래가격 보기'를 참조하여 해당하는 번호를 입력해주시십시오.

구분		대기업	중소기업	정부·공공	개인	합계
구매처	비중	%	%	%	%	100%
	거래가격(번호기입)					
판매처	비중	%	%	%	%	100%
	거래가격(번호기입)					

▣ 거래가격 보기

① 완전 무료 ② 광고/홍보 조건 무료 ③ 파격 할인된 가격 ④ 적절한 가격 ⑤ 기타()

문13. [구매 경험자 대상] 귀사가 데이터를 거래하는 주된 목적은 무엇입니까?

1순위		-	2순위	
-----	--	---	-----	--

- ① 자산·상품으로서 데이터 거래 ② 데이터서비스의 원천데이터로 활용
③ 비즈니스 전략 수립을 위한 데이터 분석 ④ 마케팅 자료로 활용
⑤ 데이터 분석 교육 및 솔루션 개발 테스트 ⑥ 데이터 기반 신규 비즈니스 개발
⑦ 기타()

문14. [구매 경험자 대상] 귀사가 데이터 거래 시 겪는 애로사항은 무엇입니까?

1순위		-	2순위	
-----	--	---	-----	--

- ① 불합리한 데이터 거래 가격 ② 데이터 거래 절차 및 방법 미숙
③ 데이터 유통 채널 부재 ④ 데이터의 품질 유지
⑤ 개인정보 처리 기술력 및 예산 부족 ⑥ 개인정보 (일부)포함 데이터 유통·활용 시 법적 문제
⑦ 기타()

문15. 귀사가 데이터 유통·활용 비즈니스 활성화를 위해 가장 필요한 지원 사항은 무엇입니까?

- ① 데이터 구입비 지원 ② 보유 데이터의 적정한 가치 평가 ③ 데이터 유통·홍보 채널
④ 데이터 유통 관련 법·기술 교육 ⑤ 기타 ()

문16. 귀사가 개인정보가 포함된 데이터를 유통·활용하는 경우 가장 필요한 지원 사항은 무엇입니까?

- ① 개인정보 등 데이터 유통·활용 관련 법제도 개선 ② 개인정보 비식별화 등 처리기술 지원
③ 개인정보 활용에 대한 인식 개선 ④ 개인정보를 유통하는 별도의 플랫폼 개발
⑤ 기타 ()

※ 지금부터는 전체 응답 질문입니다.

V. IT 및 R&D 투자 관련 질문

문17. 귀사의 데이터 관련 사업 추진을 위한 투자 비용은 매출 대비 몇 %인지 기입해주십시오.

구분	2016년	2017년 예상(전체)
데이터 관련 사업 투자 비중	(2015년 매출 중 몇 % 투자) %	(2016년 매출 중 몇 % 투자) %

문18. 귀사에서 데이터 관련 사업 추진을 위한 세부 투자금액(예산) 비중을 기입해주시십시오.

구분	2016년	2017년 예상(전체)
① 설비 투자	%	%
② R&D 투자	%	%
③ 라이선스 지출	%	%
④ 마케팅/홍보	%	%
⑤ 해외 투자	%	%
⑥ 기술교육/훈련 투자	%	%
⑦ 기타 :	%	%
전체	100%	100%

☐ 투자금액(예산) 항목 설명

구분	설명
① 설비 투자	솔루션 개발, 서비스 운영 등을 위한 인프라 투자 비용
② R&D 투자	솔루션, 서비스 등을 기획, 개발을 위한 연구를 위해 투자한 비용
③ 라이선스 지출	솔루션 개발, 서비스 판매 등을 위해 한 해 동안 지불 되는 라이선스 비용
④ 마케팅/홍보	마케팅/홍보를 직접 하거나 타 업체에 위탁하는 경우 소요 비용
⑤ 해외 투자	해외 진출을 위해 투자한 비용
⑥ 기술교육/훈련	직원에게 관련 기술 및 지식 교육을 위해 투자한 비용
⑦ 기타	그 외 데이터 사업을 위해 투자한 비용

VI. 시장 환경에 관한 질문

문19. 귀사의 데이터 관련 주력 사업 분야에서 **주요 고객**은 누구인지 체크해주시요.(한개만 응답)

- ① 국내 기업 고객 ② 국내 개인 고객 ③ 국내 정부/공공
④ 해외 기업 고객 ⑤ 해외 개인 고객 ⑥ 해외 정부/공공
⑦ 기타 : _____

문20. 귀사의 데이터 관련 주력 사업 분야에서 **가장 주된 경쟁국**은 어느 나라이며, **경쟁사**는 어떤 기업인지 기입해 주십시오.(한개만 응답)

- ① 미국 (기업명 : _____) ② 중국 (기업명 : _____)
③ 유럽 (기업명 : _____) ④ 일본 (기업명 : _____)
⑤ 한국 (기업명 : _____) ⑥ 기타 (기업명 : _____)

VII. 데이터 및 빅데이터 산업 활성화를 위한 의견수렴

문21. 귀사가 데이터 관련 사업 수행시 **가장 큰 애로사항 2개**를 선택해주시요.

1순위		-	2순위	
-----	--	---	-----	--

- ① 시장 포화 ② 동종업계 내 경쟁심화(가격덤핑, 원백 등)
③ 고정비용(임대료, 인건비 등) 상승 ④ 기술개발 부담 등 자금난
⑤ 데이터 전문 인력 부족 ⑥ 수익모델 부재
⑦ 신규 기술개발(R&D) 역량 미흡 ⑧ 유료화 인식 부족(데이터이용료, 유지보수료, 제품가격 등)
⑨ 내수시장 악화 ⑩ 국제경기 침체
⑪ 해외 판로 개척 어려움 ⑫ 법·제도적 제약
⑬ 데이터 표준화 미비 ⑭ 기타()

문22. 귀사의 데이터 관련 사업 추진 및 산업 활성화를 위해 **가장 필요한 정책적 지원 사항 2개**를 선택해주시요.

1순위		-	2순위	
-----	--	---	-----	--

- ① 해외 진출 및 신규 판로 개척 지원 ② 기술개발, 기술이전 등을 위한 예산(자금) 지원
③ 세제 혜택 지원 ④ 컨설팅 및 품질관리 등 기술적 지원
⑤ 전문인력 양성 및 교육 지원 ⑥ 시장·기술 동향 정보 제공
⑦ 관련 법·제도 개선 ⑧ 공공·민간의 역할 분담
⑨ 데이터 표준화 지원 ⑩ 기타()

문23. 귀사에서 데이터 관련 사업 수행 시 **가장 필요로 하는 정보 2개**를 선택해주시요.

1순위		-	2순위	
-----	--	---	-----	--

- ① 데이터 관련 시장 정보(산업통계, 동향분석정보, IT트렌드 등)
② 데이터 관련 전문인력 정보(DBA, 기획전문가, 데이터 분석가 등 인력풀, 채용정보 등)
③ 데이터 관련 제품/서비스 정보(SW, 서비스 소개정보, 데이터 거래·유통정보 등)
④ 데이터 관련 교육 정보(세미나, 교육과정, 자료집 등)
⑤ 데이터 관련 정부지원사업 정보(제작지원, 사업화지원, 해외진출지원 등)
⑥ 데이터 관련 해외 시장 정보(해외 시장 동향 등)
⑦ 데이터 관련 국내외 표준 정보(데이터 형식, 데이터 관리표준인증 등)
⑧ 기타()

문24. [문1 ③ 데이터 서비스 업체만 응답]

데이터서비스 사업을 하는 귀사는 국제회계기준상의 5가지 무형자산 유형 중에서 **기술 관련 자산**이 전체 매출에 기여하는 비중은 몇 %라고 생각하십니까?

 %

■ 국제회계기준 상의 무형자산 5가지

유형	예시
① 마케팅 관련 자산	상표권, 인터넷 도메인주소
② 고객 관련 자산	고객정보
③ 예술 관련 자산	문학작품, 저작권
④ 기술 관련 자산	특허기술, 데이터베이스
⑤ 계약 관련 자산	서비스 관련 계약

문25. [문1 ③ 데이터 서비스 업체만 응답]

기술 관련 자산의 세부 유형 중에서 **데이터베이스가** 전체 매출에 기여하는 비중은 몇 %라고 생각하십니까?

 %

■ 기술 관련 자산의 세부 유형

유형
① 특허기술
② 컴퓨터 소프트웨어 및 배치설계(maskworks)
③ 비특허기술(노하우)
④ 데이터베이스
⑤ 기밀사항(비밀공식, 프로세스 또는 조리법과 같은 영업비밀)

문26. 데이터 산업 활성화를 위한 개선사항, 건의사항, 기타 의견이 있으시면 구체적으로 기술해 주십시오.

♣ 응답해주셔서 대단히 감사드립니다. ♣

부록1

2017년 데이터산업분류 세부시장 정의

□ 데이터산업 범위

구분	설명	제외영역
데이터 솔루션	- DBMS, DBMS관리, 데이터 모델링, 분석·시각화, 검색엔진, 품질 등 관련 솔루션 제품으로 비즈니스를 영위하는 기업 - 라이선스, 유지보수, 커스터마이징(개발)에서 매출 발생	한글, OS 등 오피스 및 컴퓨터 운영프로그램, 게임, 이미지 뷰어/랜더링 프로그램, 유틸리티, 영상/음향 재생 프로그램 등 제외
데이터 서비스	데이터를 활용해 정보제공, 데이터 거래, 분석 결과 정보 등을 온·오프라인(모바일, 앱 등 포함)으로 제공하면서 데이터 이용료/수수료 또는 광고료 등으로 비즈니스를 영위하거나, 마케팅을 목적으로 데이터를 수집·가공하여 판매하는 기업(데이터 브로커)	정보전달을 주목적으로 하며, 온라인게임 등 오락적 목적이나 전자상거래의 일반재화 판매매출 등은 제외
데이터 구축	DB설계, 데이터 이행 등을 포함한 DB시스템 구축, 문서·영상·영상 등의 데이터를 DB로 변환·정비하는 데이터 처리, 데이터 외부 제공을 위한 API, LOD 구축, DW, Data Lake 등의 데이터 구축 등의 비즈니스를 영위하는 기업(데이터/DB 관련 SI·IT아웃소싱 포함)	네트워크 시스템 구축, H/W 유통 및 전산실 구축, IT아웃소싱(유지보수, 인력파견, 시스템운영), 공장/업무자동화 시스템 구축, 홈페이지 구축 등 제외
데이터 컨설팅	데이터 거버넌스, 품질, 데이터 설계 및 데이터 활용 등 데이터 관련 기획 및 컨설팅 비즈니스를 영위하는 기업	교육, 일반 경영컨설팅 제외
데이터 인프라	데이터의 저장과 활용을 목적으로 하는 서버, 스토리지, 네트워크 등 기반 구조를 제조, 판매하거나 기획, 구축, 운영 또는 관련 컨설팅 비즈니스를 영위하는 기업	DBMS가 탑재되지 않는 서버와 데이터 인프라 관련 시설물 등은 제외

□ 데이터 솔루션 시장 정의

대분류	중분류	소분류	정의	관련 예시
데이터 솔루션	데이터 수집	데이터 검색	- 사용자가 원하는 데이터를 찾을 수 있도록 지원하는 솔루션 - 사용자가 필요로 하는 데이터를 다양한 검색조건을 지정하여 데이터세트를 만들고 활용할 수 있도록 지원하는 솔루션	검색엔진, Query
		로그데이터 수집	다양한 시스템으로부터 발생하는 로그를 실시간으로 수집하고, 이종간 발생하는 다양한 정보의 형태를 정형화하여 통합 관리하는 솔루션	공정로그 수집 솔루션, 고객의견(VOC) 로그 수집 솔루션
		웹데이터 수집	조직적, 자동화된 방법으로 인터넷 상에 존재하는 웹 문서들을 추적하여 필요한 데이터를 수집하는 솔루션	웹 크롤러
	데이터 설계	데이터 아키텍처	기업 또는 기관의 전사 차원에서 데이터 현황을 파악하고 관리할 수 있는 솔루션	데이터 참조모델 생성 및 활용, 다중DBMS 저장소, 데이터 아키텍처 관리 솔루션 등
		데이터 모델링	데이터베이스 구축을 위해 데이터 모델링을 수행하고 모델링 결과에 따라 데이터베이스 물리설계도를 생성하는 솔루션	데이터 모델링 솔루션 등
	DBMS	디스크 DBMS	기존에 주로 이용하고 있는 DBMS로서, 데이터를 별도의 물리적 저장장치에 보관하는 DBMS	전통적 DBMS
		인메모리 DBMS	새롭게 등장한 DBMS로서, 활용하고자 하는 데이터를 인메모리(in-memory)에 올려놓고 운영함으로써 성능이 개선된 DBMS	실시간 DBMS, 인메모리 기반 DBMS 등
		DBMS 어플라이언스	DBMS가 운영될 하드웨어와 저장장치를 DBMS와 통합하여 하나의 하드웨어로 제공하는 솔루션 또는 서버	DBMS가 내재화된 서버
		기타 DBMS	전통적인 DBMS, 인메모리 DBMS, DBMS 어플라이언스에 포함되지 않는 DBMS	클라우드 DBMS, 하이브리드 DBMS, 기타DBMS 등
	데이터 관리	DB 운영 관리	- DB 운영 단계에서 적용하는 도구로서, 주로 DB 형상관리(configuration management)에 적용되는 솔루션 - 비즈니스 연속성을 위한 데이터의 관리를 위한 백업/복구 소프트웨어	DB 운영 솔루션, 백업 및 복구, 데이터 중복 제거 등

대분류	중분류	소분류	정의	관련 예시
데이터 솔루션	데이터 관리 (계속)	DB 성능 관리	• DB성능을 모니터링하고 장애 및 지연 요인을 발견하고 해결하기 위한 솔루션	대용량 데이터 처리, 자동진단(장애분석), 자동튜닝, 능력(CAPA) 산정, DB성능 모니터링 솔루션
		DB 보안 관리	• DB의 유출, 개량, 파괴, 접근 등으로부터 DB의 보안을 위한 방화벽, 침입탐지/방지, 서버 보안 등 데이터베이스 접근 통제, 암호화, 보안감사를 수행하는 소프트웨어	멀티DB 암호화 및 접근 통제 연계, 클라우드DB 보안, 능동적DB보안, 대규모 스트림 데이터 암호화, 데이터 익명화, DB포렌식(Forensic) 감사 등 DB 암호화 및 접근통제 솔루션
		데이터 분산 처리	• 기존의 데이터 처리 기술로는 효율적으로 처리하기 어려운 대용량의 비구조화 데이터 처리를 위하여 대량의 하드웨어를 이용하여 분산 저장(분산파일시스템, 분산 스토리지 시스템), 처리(맵리듀스)하는 솔루션	클라우드 컴퓨팅, 하둡
		데이터 통합·흐름 관리	• 운영 데이터베이스로부터 데이터를 추출하여 데이터웨어하우스로 정제 및 가공하여 적재하는 솔루션, 기업 및 기관 내부의 데이터 흐름을 자동화하는 솔루션	EAI(Enterprise Application Interface), ETL(Extract, Transform, Load)
데이터 품질	데이터 품질 관리	데이터 품질 관리	• 운영 데이터베이스에 보관되고 있는 데이터의 품질을 측정하고 평가하며, 변경 영향을 분석하는 솔루션	실시간 데이터 품질 검증, 비정형데이터 쿼리 매칭 등 데이터 품질 관리 솔루션
		메타데이터 관리	• 운영 데이터베이스의 데이터 정의와 표준 등 메타데이터를 관리하고 데이터 변경을 통제하는 솔루션	데이터 표준 자동화, 대용량 분산 메타데이터 관리, 영향도 분석 등 메타데이터 관리 솔루션
		마스터데이터 관리	• 기업 활동의 기준이 되는 핵심 데이터를 식별하고 전사적으로 일관되게 사용하기 위해 적용하는 솔루션	플랫폼형 마스터데이터 관리 기술, 개발 프레임워크 일체형 마스터데이터 관리 도구, MDM과 빅데이터 결합, 마스터데이터 추출 및 분리 기술 등 Master Data Management Solution
데이터 분석	정형 데이터 분석	정형 데이터 분석	• DBMS의 데이터를 사용자의 이용목적에 맞도록 간편하게 보고서를 생성 및 활용 할 수 있도록 지원하는 솔루션(데이터리포팅) • 미리 정의된 관점에 따라 사용자가 기준을 유연하게 변경하면서 데이터를 분석할 수 있도록 지원하는 솔루션(OLAP) • 데이터 집합을 통계적으로 분석하여 의미 있는 인사이트를 찾아내는 과정을 지원하는 솔루션 등(데이터마이닝)	Reporting 솔루션, OLAP(Online Analytical Processing), 통계분석패키지
		비정형 데이터 분석	• 웹사이트 데이터를 추출 및 분석하여 기술트렌드 및 기업 평판 등을 분석하기 위해 적용하는 솔루션(웹크롤링) • 주로 문서 및 웹사이트의 데이터에서 키워드를 찾아내고 빈도 및 연관관계 등을 분석하는 솔루션 등(텍스트마이닝) • 신문기사, 논문, 특허 등의 동향정보로부터 주요 개체 및 개체와 관련된 사실정보를 정형화된 형태로 추출하여 분석에 활용하는 솔루션(정보추출) • 음성 및 영상 등의 데이터로부터 의미 있는 데이터를 추출하여 분석하는 솔루션 • 소셜미디어 내의 인맥정보를 이용하여 사회관계망을 분석하고, 포스팅 되는 내용을 통해 최신 이슈와 트렌드를 분석하는 솔루션	웹 크롤링(Web Crawling), 텍스트 마이닝, 비디오데이터 가공 솔루션, 음성 데이터 가공 솔루션, 트위터, 페이스북 등의 API 연계 솔루션(Social Analytics 솔루션)
	실시간 데이터 분석	실시간 데이터 분석	• 데이터 분석을 통해 위험 신호 또는 징후를 감지하여 위기나 재난에 대처할 수 있도록 예측하는 솔루션	CEP(Complex Event Processing), BAM(Business Activity Monitoring)
		데이터 시각화 분석	• 데이터 세트의 의미를 시각적으로 표현하거나 탐색하는 과정을 지원하는 솔루션	시각화 및 비주얼 분석 솔루션(Visual Analytics 솔루션)
	데이터 플랫폼	데이터 플랫폼	• 개인, 사업자, 기관 등이 보유한 원천데이터를 변환, 가공, 정제, 발행(publishing), 저장, 질의 등이 가능하도록 하는 솔루션 혹은 기능들의 집합	빅데이터 플랫폼 등

□ 데이터구축/컨설팅 시장 정의

대분류	중분류	소분류	정의	관련 예시
데이터 구축/ 컨설팅	구축	DB설계·구축	- 업무요건을 충족시킬 수 있는 데이터베이스의 구조와 형태 및 속성을 정의하는 서비스 - 정형 또는 비정형의 대량 데이터를 가공전 raw data로 저장하는 리포지토리나 사용자가 쉽게 접근할 수 있는 분석용(의사결정) 데이터베이스 등을 구축하는 서비스	물리 DB 설계, 데이터웨어하우스(DW), 데이터마트(Data Mart), 데이터레이크(Data Lake) 등 구축
		데이터 이행·처리	- 시스템 재구축으로 인해 기존에 운영 중인 데이터베이스에 보관 중인 데이터를 점검하고 정제하여 새롭게 개발하고 있는 데이터베이스로 이행하는 서비스 - 일반문서, 음성, 영상 등의 자료를 데이터베이스에 보관 할 수 있도록 형태를 정비하거나 변환하는 서비스 - 다양한 데이터를 주어진 형식에 맞추어 구축하는 서비스	차세대 시스템 구축, 데이터 변환 등
		기계처리형 데이터 구축	- 기업 또는 기관이 보유하고 있는 데이터를 외부에 제공 및 활용할 수 있도록 API, LOD(Linked Open Data) 등을 개발, 구축하는 서비스 - 기존에 존재하지 않았던 데이터를 특별 목적을 위해 데이터화(datafication) 하는 것 포함	API 구축, LOD 구축 등 machine processable 데이터 구축
	컨설팅	데이터 설계	- 전사 차원에서 데이터 요건을 분석하여, 단위 데이터 저장소를 정의하고 데이터 저장소들의 연관 관계를 정의하는 서비스 - 데이터 아키텍처를 기반으로 특정 영역의 업무 요건을 보다 상세하고 완전하게 분석하여 데이터베이스를 구축할 수 있는 수준으로 설계하는 서비스 - 데이터 설계를 기획하거나 가이드, 리딩 또는 지원하는 서비스	Data Architecture Planning, Analytics Strategy Planning, Data Modeling, Data Reference Model 등
		데이터 품질	운영하고 있는 데이터베이스에 포함된 실제 데이터의 품질을 점검하고, 개선점을 제시하는 서비스. 데이터 정의, 데이터 값, 데이터 연관관계 등의 품질 영역이 있음	데이터 품질 심사 및 인증, 데이터 품질 진단 등
		데이터 관리 성능개선	운영하고 있는 데이터베이스의 활용 과정에서 온라인 및 배치 작업의 병목 지점을 발견하고 이를 개선함으로써 요구하는 서비스 수준을 달성시키기 위한 서비스	DB튜닝, DB성능 진단 등
		데이터 거버넌스 체계	데이터 아키텍처로부터 데이터 모델, 물리 데이터베이스 설계안 등이 현재 운용 중인 데이터베이스의 모습과 동시성을 유지하면서, 변경 요인을 효과적으로 반영할 수 있는 조직, 역할, 프로세스 등을 설계하는 서비스	데이터 거버넌스 진단, 데이터 거버넌스 체계 수립 등
		데이터 분석	내부 및 외부 데이터를 활용하여, 기업 또는 기관에서 필요로 하는 분석 기법을 개발하고, 테스트하여, 분석 모듈을 개발할 수 있도록 컨설팅해주는 서비스	추천엔진 개발, FDS(Fraud Detection System) 개발 등

□ 데이터서비스 시장 정의

대분류	중분류	소분류	정의	관련 예시
데이터 서비스	데이터 거래	데이터 마켓	• (데이터 중개) 데이터 유통 플랫폼을 통해 데이터 보유자(보유기업)와 수요자(수요기업)간 원천데이터를 중개하는 서비스 • (데이터 판매) 온오프라인, API, 자사 유통시스템 등을 이용해 기업간(B2B)에 원천데이터를 판매하는 서비스	API스토어, 데이터허브, 데이터스토어, 데이터 오픈마켓, 데이터중개 등
		데이터 신디케이션	• 다양한 데이터를 바탕으로 고객이 원하는 형태로 또는 특정 주제에 따라 구성하여 제공하는 맞춤형 데이터 서비스	데이터 융합
		교육/훈련	• 자격증, 시험, 유학, 연수 등의 교육/훈련 관련 정보서비스	
	정보 제공	신용/재무	• 기업일반, 기업·개인 신용, 재무·회계, 보험/증권/주식/환율 등의 정보서비스	
		통계	• 시장정보, 리서치정보 등 통계, 시장동향 정보서비스	
		경영/비즈니스	• 쇼핑·고객, 제품·상품·부품, 물류·운송, 경매·입찰, 부동산 정보서비스	
		취업/창업	• 취업(구인구직), 창업 관련 정보서비스	
		뉴스	• 뉴스제공, 뉴스사진제공 등 뉴스를 제공하는 서비스	
		포털	• 다양한 정보를 검색해 접근할 수 있도록 하는 포털 형식의 정보서비스	
		행정/법률	• 법률·법규·판례, 행정·민원, 조세·세법, 지식재산권 관련 정보서비스	
		학술	• 인문사회과학, 자연과학, 기술과학, 의학, 참고자료 등 관련분야 전문정보 및 논문 정보서비스	
		문화/예술	• 문학·만화, 미술, 사진, 음악, 공연예술, 매체예술, 전통문화 등 문화예술 관련 정보서비스	
		생활	• 관광·여행, 건강·보건, 교통, 지도·지리, 날씨·기상, 인물, 도서·출판, 종교, 결혼·장례, 여성·육아, 오락·게임, 스포츠·레저 등 생활과 밀접한 정보를 제공하는 정보서비스	
	데이터 분석 제공	환경분석 서비스	• 소셜데이터, 오픈데이터 등 외부데이터와 자체보유하고 있는 데이터를 연계하여 주요 환경 변화 트렌드를 정보로 생성하여 제공하는 서비스	Horizontal Scanning, Reference Engine, 부정사용방지 또는 이상거래 탐지시스템(FDS; Fraud Detection System), SNS 대화 감성분석 서비스, 여론분석서비스, 상권분석서비스 등
		예측분석 서비스	• 서비스 및 상품추천, 위험 사전 예보 등의 분석을 대행하여 수행하고 그 결과를 제공하는 서비스	
		심사분석 서비스	• 개인신용등급 등 금융 및 상거래를 위해 필요한 심사 분석을 대행하여 수행하고 그 결과를 제공하는 서비스 등	
		기타 데이터분석 서비스	• 그외 대량의 다양한 데이터를 분석하여 유의미한 값을 도출, 제공하는 서비스	

■ 데이터 서비스 - 데이터 판매와 정보제공 구분

※ 최종이용자에게 정보전달을 목적으로 데이터를 가공해 무료 또는 서비스이용료를 부과해 제공하는 서비스는 “정보서비스”입니다. B2B로 원천데이터를 판매하는 경우에는 “데이터거래/데이터판매”입니다.

(예: 학술논문검색 및 원문다운로드 = 학술정보서비스 / 학술논문 목록데이터 판매 = 데이터판매)

□ 데이터인프라 시장 정의

대분류	중분류	소분류	정의	관련 예시
데이터 인프라	온프레미스	서버 (DB서버)	- DBMS를 설치하여 데이터 저장 및 활용에 사용되는 H/W - 데이터 구축을 통해 확보된 데이터를 활용하거나 서비스하기 위한 서버 기능을 제공하는 H/W	Stand Alone 서버, 이중화 서버, 병렬 클러스터 등
		스토리지	- 데이터 저장하고 저장된 데이터의 효과적인 공유를 위한 기능과 용량을 갖춘 전용 장치 - 데이터 구축을 통해 확보된 데이터를 저장하고 안정적으로 유지, 관리하는 기능을 갖춘 H/W - 유사시에 대비하여 데이터의 복사본을 저장하도록 기능과 용량을 갖춘 별도의 H/W	디스크 스토리지, SSD, Flash Disk, NAS
		네트워크	- 데이터 저장 및 활용을 위한 DB서버와 스토리지가 포함된 유·무선 네트워크 - 데이터의 백업/복구, 재해복구를 위한 원격지 데이터센터를 연결하는 전용 네트워크	백업 및 복구 전용망, 재해복구 데이터 센터 전용망
	클라우드	클라우드 구축	클라우드 서비스 플랫폼을 구축	네이버 클라우드 플랫폼, iBot
		클라우드 서비스 제공	- 클라우드 스토리지를 제공 - 클라우드 서비스 소비자와 제공자 사이에서 클라우드 서비스의 부가가치 창출을 위해 소비자를 대신하여 일하는 중개자로, 소비자와 제공자간 관계 조율 및 소비자의 요구에 맞춰 최적의 클라우드 서비스를 제안하고 다양한 클라우드 서비스의 활용, 성능 관리, 전달 등을 담당	Dropbox, Google Drive 등과 같은 유·무료 클라우드 스토리지를 제공하는 IaaS(Infrastructure as a Service) 유형의 클라우드 서비스

The background features a stylized city skyline in shades of blue and grey on the left side. The rest of the page is composed of large, overlapping geometric shapes in various shades of blue, creating a modern, abstract design.

부록2.

국내 주요 데이터기업

□ 국내 주요 데이터기업

기업명	주 사업내용
광개토연구소	- 특허평가데이터, US특허거래데이터, US특허소송데이터, 특허관리전문회사(NPE) 데이터 판매 및 분석서비스 - 주요제품 : Patentpia.com
그로투	- 데이터 통합 솔루션 - 주요제품 : BTL Data Integrator
누리미디어	- 국내 학술지 논문 서지데이터 판매 및 정보제공서비스 - 주요제품 : DBpia, KRpia
다음소프트	- 언어처리, 텍스트마이닝, 소셜미디어 분석, 검색 솔루션 - 주요제품 : Mining Search
다이퀘스트	- 데이터 수집, 검색, 언어처리 솔루션 - 주요제품 : Mariner 4, Diver 5, i-Spider 3, DISA, DQ-Cat, DQ-Clu, DRAMA
데이터스트림즈	- 데이터 흐름관리, 데이터 품질관리, 메타데이터관리, 마스터데이터관리, 데이터 플랫폼 솔루션 - 주요제품 : TeraStream, Terastream for Hadoop, MetaStream, QualityStream, FACT, TeraTDS, TeraNRT, DeltaStream, Q-Track, ImpactStream, MasterStream, TeraStream BASS
리얼타임테크	- DBMS - 주요제품 : Kairos-RDBMS, Karios-Spatial, Kairos-MO, Kairos-LITE, RTSync
부동산114	- 아파트 단지별 데이터, 아파트 시세 데이터, 아파트 분양권 시세 및 단지데이터, 재건축 단지 데이터, 아파트 투자지표 데이터 판매 및 정보제공서비스 - 주요제품 : 부동산114
비아이매트릭스	- 비즈니스인텔리전스, 분석리포팅, 비정형 분석 솔루션 - 주요제품 : MATRIX, i-CANVAS, i-BIG
비투엔	데이터 모델링, 빅데이터 아키텍처, 빅데이터 거버넌스, 빅데이터 분석 및 시각화, 데이터 아키텍처, 데이터 품질 관리, BI 및 데이터웨어하우스, 데이터 이행, 데이터 성능개선 등 데이터 컨설팅 및 구축
사이람	- 데이터 분석 솔루션 - 주요제품 : NetMiner, NetMetrica, NetExplorer, Sopion, NetViz
사이버다임	- 데이터 수집 및 가공, 데이터 분석, 시각화 솔루션 - 주요제품 : Process Analyzer
선재소프트	- DBMS - 주요제품 : SUNDB, GOLDILOCKS
솔트룩스	- 실시간 빅데이터분석, 시맨틱 검색, 비정형 마이닝, 데이터 수집, 데이터 시각화, 자연어처리 솔루션 - 주요제품 : BigO, DISCOVERY, TORNADO, BLUEBOLT, STORM, RAINBOW, LEA
씨에스리	데이터모델링, 데이터아키텍처, 데이터 품질진단, 데이터 관리체계 개선, 데이터 운영관리 및 성능개선 등 데이터 컨설팅
알티베이스	- DBMS - 주요제품 : Altibase HDB, Altibase XDB
에프앤가이드	- 주식/경제/재무/파생데이터 판매, 펀드 수익률/위험등급 등 펀드 분석서비스, 채무/CP/단기사채 등 채무증권 분석서비스 - 주요제품 : FnSearch, FnConsensus, FnOwnership, FnDB Navigator, FnScreen
엑셈	- 성능관리 솔루션 - 주요제품 : MaxGauge, InterMax
엔써티	- 데이터 백업 솔루션 - 주요제품 : BackupXcelerator, NcertiReplicator, NcertiHybrid-XR

기업명	주 사업내용
엔코아	- 데이터 아키텍처, 데이터 흐름 관리, 데이터 이행 관리, 데이터 품질관리 솔루션 - 데이터 아키텍처, 데이터 품질, 데이터 이행, 데이터 거버넌스, 데이터웨어하우스, 비즈니스 인텔리전스 컨설팅 - 주요제품 : DataWare DA#, DataWare DF#, DataWare Meta#, DataWare AP#, DataWare DQ#, DataWare ETT#
LG CNS	데이터웨어하우스, 비즈니스인텔리전스, 빅데이터 고급분석, 스마트 빅데이터 플랫폼 컨설팅
와이즈넷	- 데이터 검색, 수집, 분석, 텍스트마이닝 솔루션 - 주요제품 : Search Formula-1 V5, WISE iDESK, WISE Bot, WISE BIC Analyzer, WISE BICrawler, WISE InfoFinder, WISE TEA V2
웨더아이	- 기상 관측데이터 및 예보데이터, 일출일몰시간, 강수확률 등의 기상데이터(API) 판매 서비스 - 주요제품 : weatheri
웨어밸리	- 성능관리, 보안 솔루션 - 주요제품 : Orange, Galea, Cyclone, Chakra Max
위세아이텍	- 데이터 품질관리, 데이터 분석, 데이터 시각화 솔루션 - 주요제품 : WISE DQ, WISE OLAP, WISE Visual, WISE Advisor, WISE Meta, WISE Open, WISE Campaign
웍스	- 특허 서지/원문데이터, 상표데이터 검색 및 분석서비스 - 주요제품 : 원텔립스, 웍스온, 인투마크, WIPS GLOBAL, PATBRIDGE
이글로벌시스템	- 데이터 보안 솔루션 - 주요제품 : CubeOne
이노티움	- 데이터 백업 솔루션 - 주요제품 : LIZARD Backup v10
이씨마이너	- 데이터 분석 솔루션 - 주요제품 : ECMiner, ECMinerIMS, ECMinerLA
잡코리아	- 구인, 구직 데이터(API) 서비스 - 주요제품 : 잡코리아, 알바몬, 게임잡, 데브잡, 잡부산
지티원	- 데이터 품질관리, 메타데이터 관리 솔루션 - 주요제품 : DQMiner, MetaMiner, DataProfiler, DLMiner
카페인모터큐브	- 차량 정비 데이터, 차량 주요 장치별 상태(진단) 데이터 판매 및 서비스 - 주요제품 : Carffeine
케이사인	- 데이터 보안 솔루션 - 주요제품 : KSignSecureDB, WizLook, WizLookFDS
케이웨더	- 기상 관측, 예보, 특보, 항공기상 데이터 판매 및 서비스 - 주요제품 : Kweather
KT NexR	- 빅데이터 플랫폼 솔루션 - 주요제품 : NDAP, Lean Stream
케이티디에스	비즈니스인텔리전스, 데이터웨어하우스, 대용량 분산 분석 및 처리 등의 빅데이터 구축 및 컨설팅
KTH	- 데이터 API 중개서비스(유통플랫폼) - 주요제품 : APIStore
코난테크놀로지	- 데이터 수집, 분석 솔루션 - 주요제품 : KONAN Social Crawler, KONAN Web Crawler, KONAN Search 4, KONAN MemChecker, KonanLink, KONAN Text Analytics, KONAN Log Analytics

기업명	주 사업내용
코리아크레딧뷰로	- 기업 및 개인 신용데이터 제공 및 리스크관리 분석서비스 - 주요제품 : AllCredit, 크레딧파트너
코리아피디에스	- 에너지, 석유화학, 철강, 비철금속, 농산물 등 국제원자재 가격 데이터, 생산량, 소비량, 재고량 등 통계데이터 판매 및 서비스 - 주요제품 : KoreaPDS
코스콤	- 주식, 선물옵션, 채권, 외환 등 금융데이터 판매 및 서비스 - 주요제품 : Data Mall, 시세정보서비스, CHECK Expert+
큐브리드	- DBMS - 주요제품 : CUBRID 9.3
토마토시스템	- 데이터 설계 솔루션 - 주요제품 : eXERD
투비웨이	- 마스터데이터관리 솔루션 - 주요제품 : TOBEway Enterprise MDM
투이컨설팅	비즈니스 인텔리전스, 데이터 아키텍처, 데이터 품질, 마스터데이터, 오픈데이터, 디지털 트랜스포메이션, 데이터 사이언스(빅데이터 분석) 등 데이터 컨설팅
티맥스소프트	- DBMS - 주요제품 : Tmax Tiberio, Tmax ZetaData, Tmax AnyMiner, Tmax DataHub, Tmax ProSync
파수닷컴	- 데이터 보안 솔루션 - 주요제품 : Fasoo Total Privacy Solution
피애피시큐어	- 데이터 보안 솔루션 - 주요제품 : DBSAFER, INFOSAFER, DATACRYPTO
한국기업데이터	- 기업 개요, 대표자, 재무, 신용, 상장 등 기업데이터 판매 - 주요제품 : 크레탑(Cretop)
한국물가정보	- 토목/조경/기계/전기/공구 등의 물가 및 적산데이터 판매 및 서비스 - 주요제품 : 온라인 KPI서비스(물가정보API), Web-Cost
한국학술정보	- 학술지 논문 서지데이터 판매 및 서비스 - 주요제품 : KISS

2017년 데이터산업 현황 조사

발행일	2018년 4월
발행처	과학기술정보통신부 www.msit.go.kr
조사수행기관	한국데이터진흥원 www.kdata.or.kr
조사실사기관	(주)마크로밀엠브레인 www.embrain.com

1. 본 보고서는 과학기술정보통신부의 출연금으로 수행한 데이터베이스산업 육성 사업의 결과입니다.
 2. 본 보고서의 내용을 발표할 때에는 반드시 「과학기술정보통신부 · 한국데이터진흥원, 2017 데이터산업 현황 조사」 결과임을 밝혀야 합니다.
 3. 본 보고서 내용과 관련한 문의는 한국데이터진흥원 정책기획실(02-3708-5367)로 연락해 주시기 바랍니다.