

データの価値計測に係るアンケート調査結果

2021年6月4日
事務局

- 昨期の「A I時代のデータ経済政策」の検討の成果を踏まえ、データの経済価値についての調査・分析の精緻化を図るため、今期も企業向けのアンケート調査を実施。
- 実施に当たっては、調査・分析の精緻化に加えて継続的なデータの把握も念頭に、次のとおり見直しを行った。

<実施面>

- 財務情報を企業活動基本調査から入手するように変更し、調査対象を15,000社まで増加。その結果、回答数は昨期の569社から今期は2,583社まで増加。
- データの価値創出メカニズムを検討するため、アンケート調査と並行してヒアリング調査（8社）を実施。

<調査票>

- 定点観測化も念頭に調査項目を見直し、データの活用体制や環境構築等を尋ねる設問等を追加。
- 企業のデータ活用の現状を把握するため、データ活用の目的・課題を回答してもらう設問を追加。

<分析面>

- データ資本に関する検討の一つとして、データの価値が年数とともにどのように減耗していくのかを観察。
- 客観指標（付加価値）だけでなく、主観的効果を用いた分析を実施。
- データ活用が効果を生むためにはどのような要素が重要なのかを明らかにするため、データの活用体制や環境構築などの重要性を検証するための分析を実施。
- 生産関数モデルについては、活用データ容量を内部入手と外部入手に分けた分析を実施。

① 企業が付加価値を生み出す生産要素として「資本」、「労働」及び「データ」を位置づけ、生産関数を推定。

$$V = A_0 K^\alpha L^\beta Data^\gamma e^{dmy}$$

ここで、Vは付加価値、Kは有形固定資産 + 無形固定資産、Lは常時従業員数、Dataはデータ変数、dmy(ダミー変数)はindmy(製造業ダミー), sizedmy(大企業ダミー)を入れて推定した。

⇒ 実証分析の結果から、活用データ容量は、他の生産要素（資本、労働）と同様に付加価値に対してプラスの関係性を持っていること、また、外部入手したデータについては、内部入手したデータに比べて貢献が大きいとの結果を得た。

データ変数	条件	n数	修正R ²	K (資本)	L (労働)	Data (データ)
活用データ容量	—	1417	0.5330	0.39 ◎	0.50 ◎	0.04 ◎
活用データ容量	$\alpha + \beta = 1$	1417	0.2554	0.39 ◎	(0.61)	0.03 ◎
内部入手した活用データ容量	—	1417	0.5327	0.39 ◎	0.50 ◎	0.03 ◎
内部入手した活用データ容量	$\alpha + \beta = 1$	1417	0.2550	0.39 ◎	(0.61)	0.02 △
1+外部入手した活用データ容量	—	1417	0.5338	0.39 ◎	0.52 ◎	0.08 ◎
1+外部入手した活用データ容量	$\alpha + \beta = 1$	1417	0.2573	0.39 ◎	(0.61)	0.08 ◎

※外部入手活用データ容量は、1+外部入手活用データ容量を変数とし、活用データ容量が0の企業は分析対象から除いた。

② データの活用体制や環境構築の状況をダミー変数として加えたモデルを推定。

⇒ 次表のモデルでは、有償で入手した外部データを活用していること及び全社的にデータ活用ができる環境を構築していることがプラスに有意（有意水準5%）となった。また、データ分析を行う専門部署の担当者が分析していること、アライアンスやコンソーシアムなど他社等を交えた共同分析をしていることがプラスに有意（有意水準10%）となった。

n数	修正R ²	K (資本)	L (労働)	Data (活用データ容量)	有償外部 データの利 用あり	無償外部 データの利 用あり	専任の責任 者がいる	兼任の責任 者がいる	全社的に データ活用 ができる環境 を構築	複数の部署 内でデータ活 用ができる環 境を構築	データ分析を 行う専門部 署の担当者	各事業部門 のデータ分 析専門の担 当者	アライアンス やコンソーシ アムなど他社 等を交えた 共同分析
1300	0.5399	0.39 ◎	0.50 ◎	0.04	0.25 ◎	0.03	-0.01	0.07	0.12 △	-0.03	0.14 △	0.04	0.33 △

(注) ◎:有意水準1%、○:有意水準5%、△:有意水準10%

全社的なデータ活用環境構築の重要性

特定の部署ごとではなく、全社的にデータを活用できる環境がある場合、データ活用の効果が高い傾向があった。IoTなどにより様々なデータの入手が可能となる中、多くの企業が保有する基本的な顧客データに加え、Web上のアクションや人の行動、機械等の動作などに関する多様かつ大量なデータをAIにより分析し、全社的に活用するメリットは今後拡大する可能性がある。

【「全社的にデータ活用ができる環境」を構築は主観効果・付加価値に対してプラス】

⇒ 取組例

- ・ データを全社的に活用することの重要性を啓発し事業戦略化を促進
- ・ データ活用のための基盤整備の促進
- ・ データの管理、情報保護、セキュリティ対策等に対する情報提供（マニュアル、ガイドライン等）の充実 等

人材育成及び組織づくりの重要性

データ活用の課題として人材、スキル・ノウハウ不足が挙げられた一方、データ分析を行う人材や専門部署が存在する企業ではデータ活用の効果が出ていた。データ活用に取り組むための人材の育成やデータ分析を専門に行う組織の構築が有効である可能性がある。

【「データ分析を行う専門部署の担当者」は主観効果・付加価値に対してプラス】

⇒ 取組例

- ・ 人材の育成方策や組織構成についての成功事例の共有、重要性の啓発
- ・ 専門人材の派遣やシェアリング等による機会の提供（主に中小企業） 等

外部連携（組織、データ）の重要性

データの活用に当たり、外部データの利用やアライアンス等による共同分析を行うことに有効性が見られた。個社が自前で入手できるデータや分析できる範囲には限界があるが、外部資源の利用により、これを補う情報や知見を得られるためと考えられ、積極的な取組が有効である可能性がある。

【「有償外部データの利用」、「アライアンスやコンソーシアムなど他社等を交えた共同分析」は主観効果・付加価値に対してプラス】

⇒ 取組例

- ・ 外部資源の活用の成功事例の共有、重要性の啓発
- ・ オープンに活用可能なデータ基盤の整備促進
- ・ データ取引市場（情報銀行、データ連携基盤）の整備促進
- ・ マッチングの機会提供によるアライアンス形成の促進 等

企業向けアンケート調査 (Web調査)

- ① 調査対象：2019年企業活動基本調査の回答企業の内、欠損値がない企業の売上高順：合計15,000社
- ② 調査時期：2021年2月19日（金）～3月19日（金）
- ③ 回収数：2,583回答（回収率17.2%）※完答数は1,828回答

業種別、規模（常時従業者数）別の回答数

業種別（日本標準産業分類）

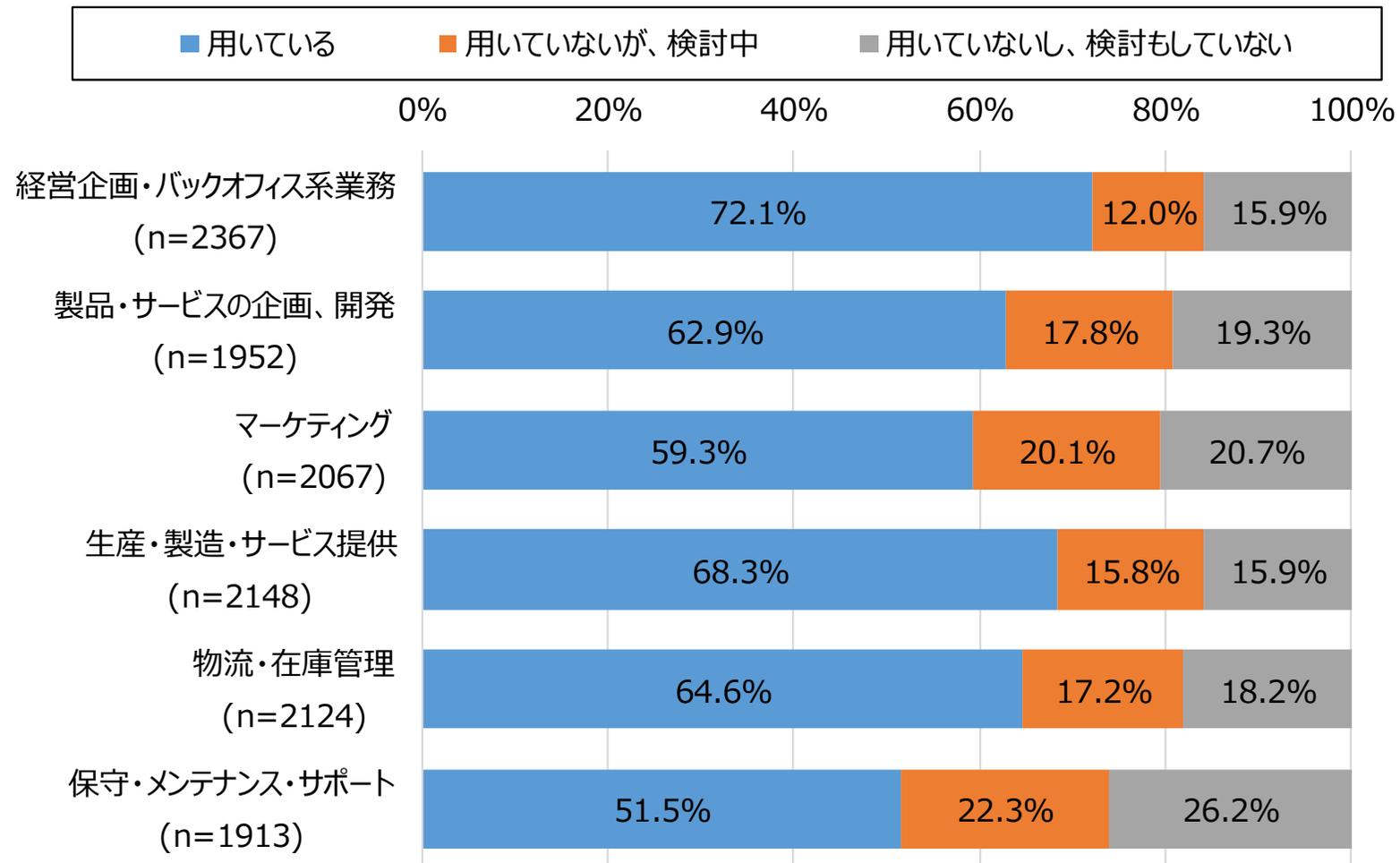
	回答数	完答数	調査対象数	回答率
農業，林業	0	0	9	0.0%
漁業	1	1	4	25.0%
鉱業，採石業，砂利採取業	1	1	14	7.1%
建設業	70	52	294	23.8%
製造業	898	634	5,783	15.5%
電気・ガス・熱供給・水道業	17	16	95	17.9%
情報通信業	120	85	869	13.8%
運輸業，郵便業	14	10	61	23.0%
卸売業，小売業	1,177	838	6,260	18.8%
金融業，保険業	10	2	47	21.3%
不動産業，物品賃貸業	59	38	254	23.2%
学術研究，専門・技術サービス業	53	32	311	17.0%
宿泊業，飲食サービス業	36	27	230	15.7%
生活関連サービス業，娯楽業	33	25	196	16.8%
教育，学習支援業	2	1	17	11.8%
医療，福祉	1	0	12	8.3%
複合サービス事業	0	0	0	-
サービス業（他に分類されないもの）	91	66	544	16.7%
計	2,583	1,828	15,000	17.2%

規模（常時従業者数）別

	回答数	完答数	調査対象数	回答率
1人～49人	1	1	17	5.9%
50人～99人	375	284	1,827	20.5%
100人～199人	690	505	3,793	18.2%
200人～299人	446	334	2,472	18.0%
300人～499人	384	259	2,427	15.8%
500人～999人	355	237	2,157	16.5%
1,000人～1,999人	169	113	1,181	14.3%
2,000人～9,999人	136	78	966	14.1%
10,000人～199,999人	27	17	160	16.9%
計	2,583	1,828	15,000	17.2%

Q1：貴社では、それぞれの業務領域において、データ分析した結果を何らかの判断に用いていますか？

※業務領域が存在する企業のみで集計

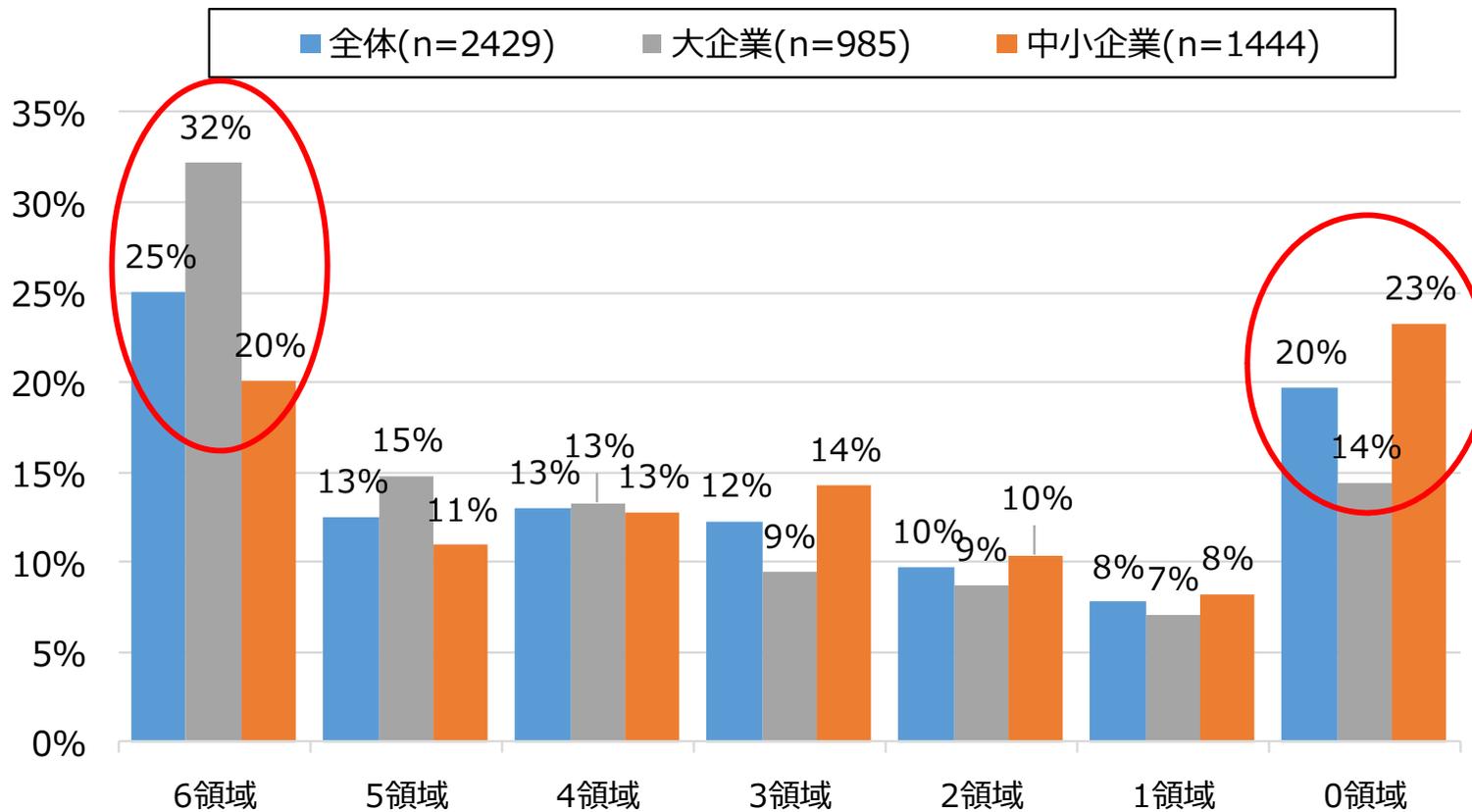


Q1：貴社では、それぞれの業務領域において、データ分析した結果を何らかの判断に用いていますか？

※データを用いている業務領域数

すべての領域（6領域）でデータを用いている企業とどの領域でもデータを用いていない（0領域）企業に二分している。中小企業では6領域よりも0領域という企業の方が多い。

データを用いている業務領域数



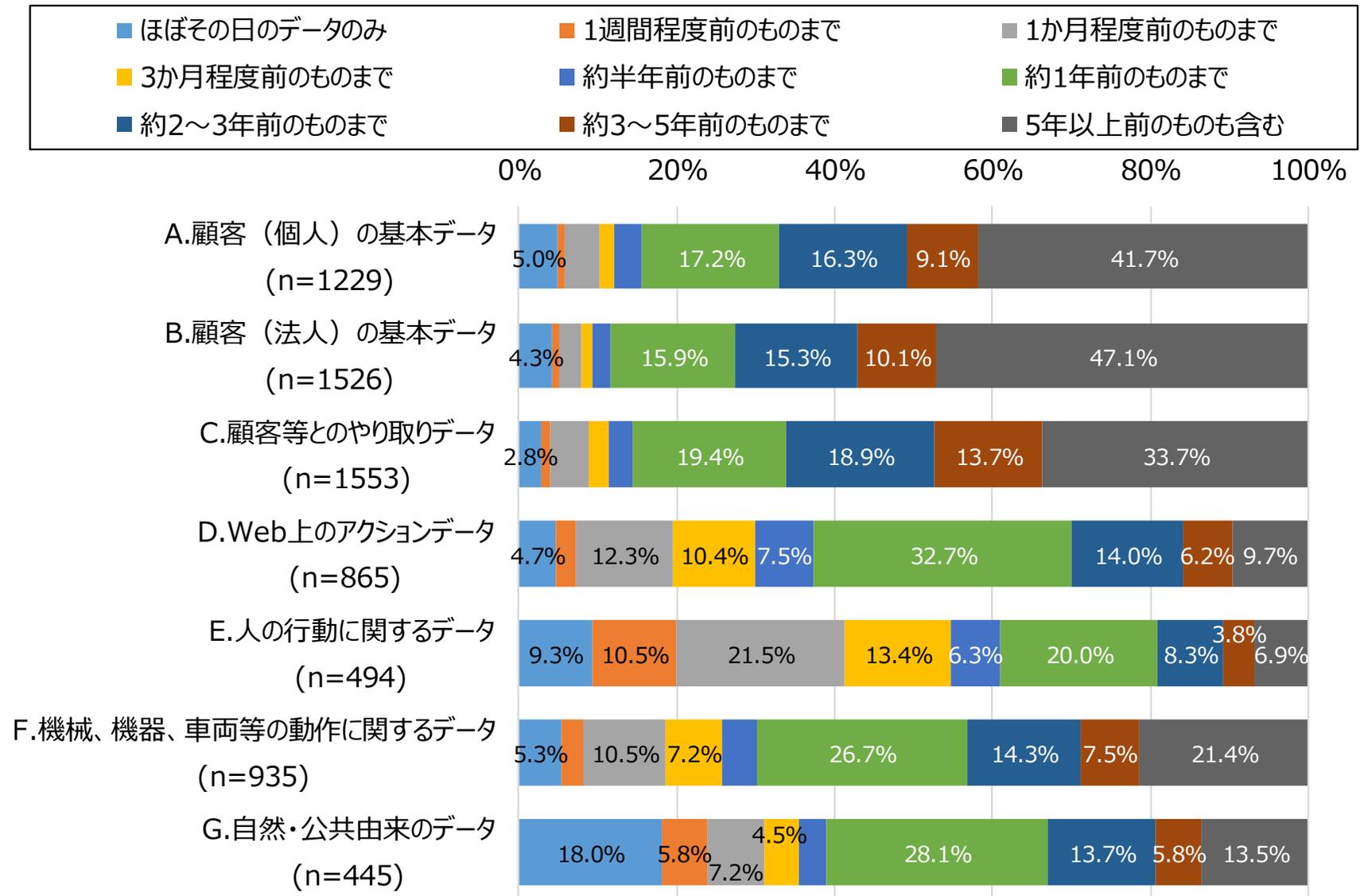
※すべての領域について無回答は集計対象から除外

データの利用状況（遡った利用）

Q4：貴社では、データ分析を行う際、おおよそどれくらい前に蓄積したデータまで遡って利用しますか？

※Q1で1つ以上選択肢1（用いている）を選択した企業のみ回答

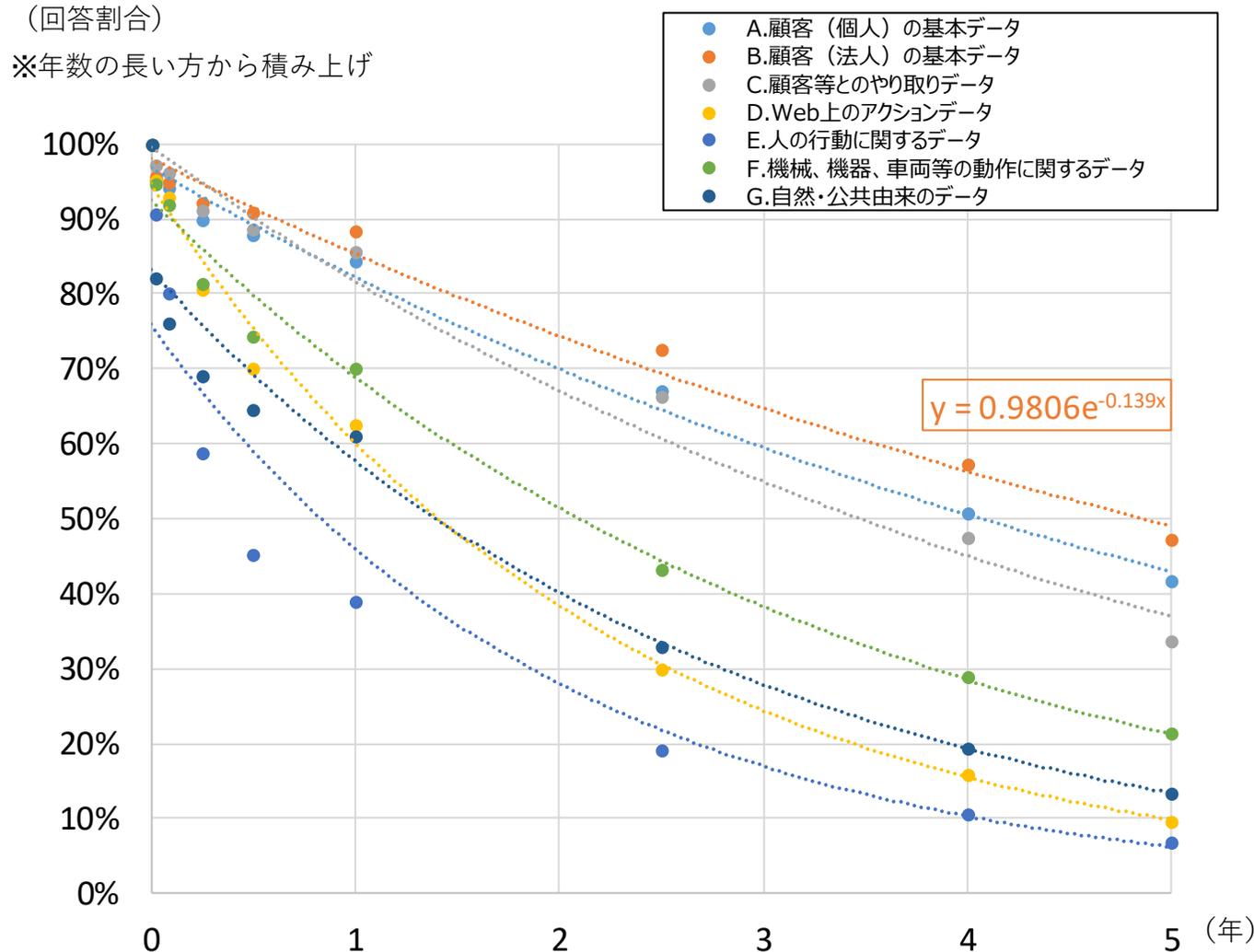
※データを蓄積している企業のみで集計。「わからない」を除いて集計。



仮に、データの価値が年数とともに減耗していくとすると、企業がどれくらい前に蓄積したデータまで遡って利用しているのか、という情報を元にデータの価値が減耗していく程度を考えることができる。

顧客の基本データは減耗が小さく、人の行動やWeb上のアクションデータは減耗が大きい結果となった。

※減耗の大きいデータについては、まだデータ活用の取組が始まったばかりであり、そもそもデータの蓄積年数が短いことも考えられ、今後継続的に観察していく必要がある。

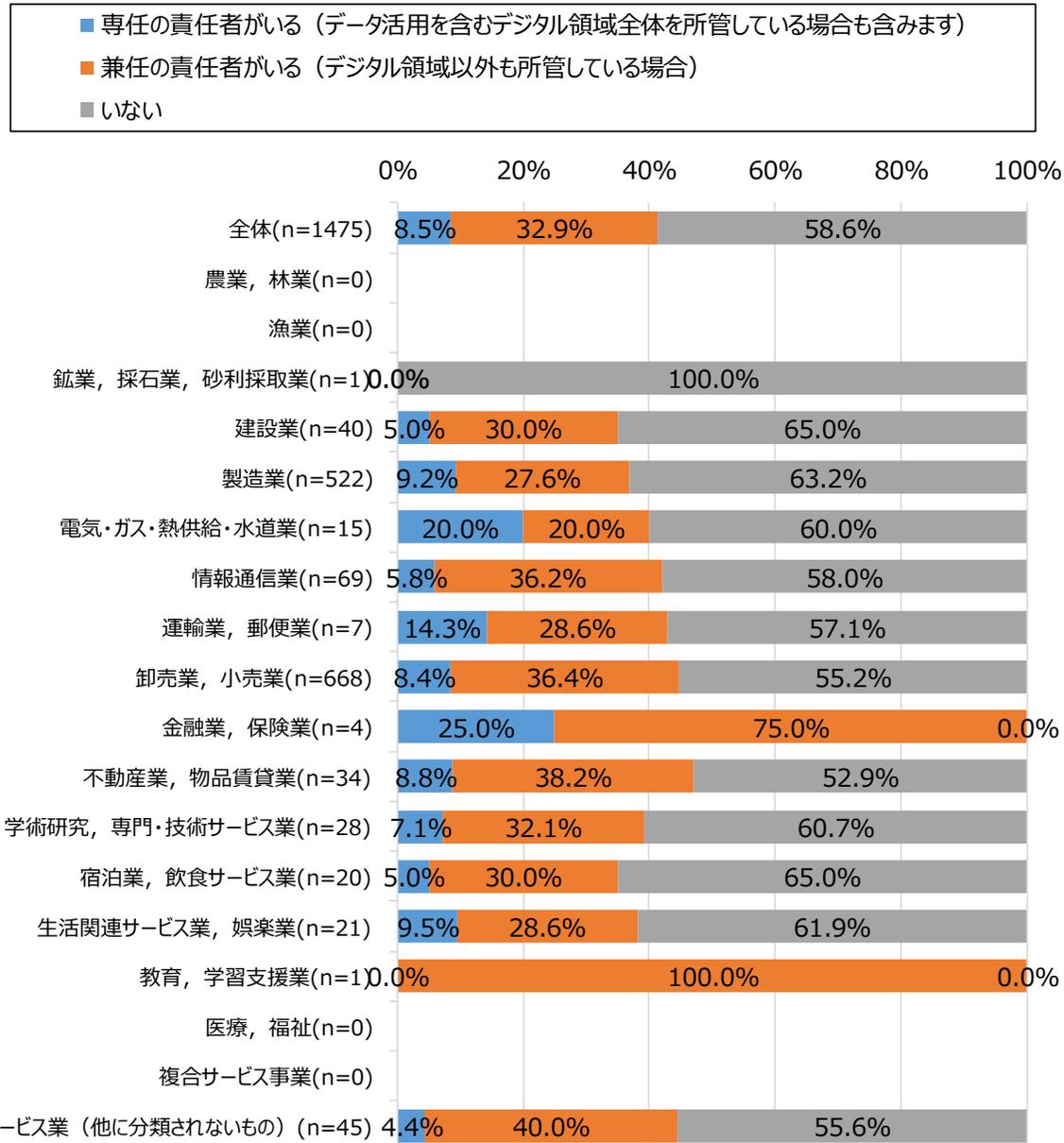


※「わからない」を除いて集計。

データ活用を主導する責任者の有無

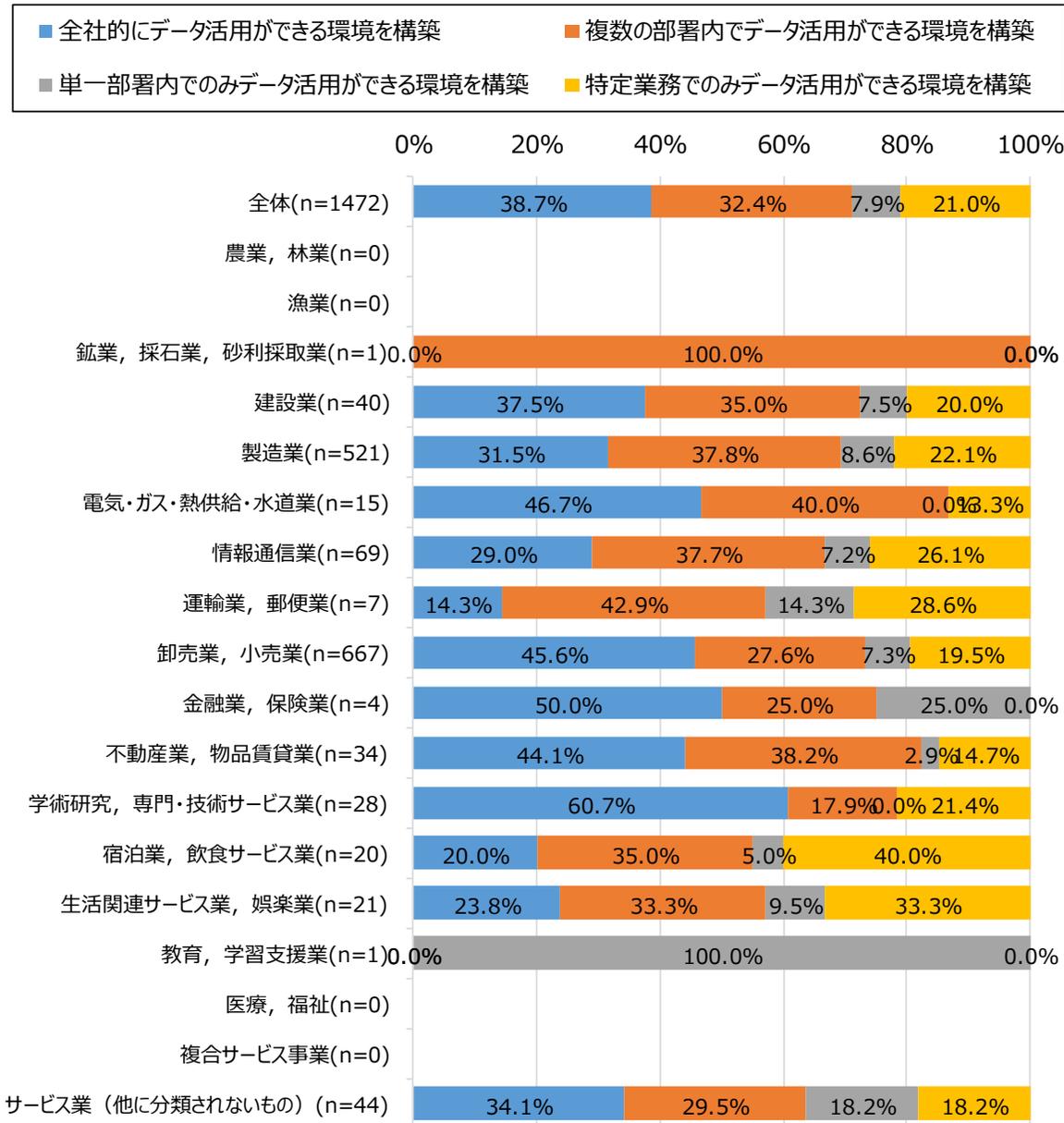
Q11：貴社では、データ活用を主導する責任者（CIO、CDOなど）はいますか？

※Q1で1つ以上選択肢1を選択した企業のみ回答



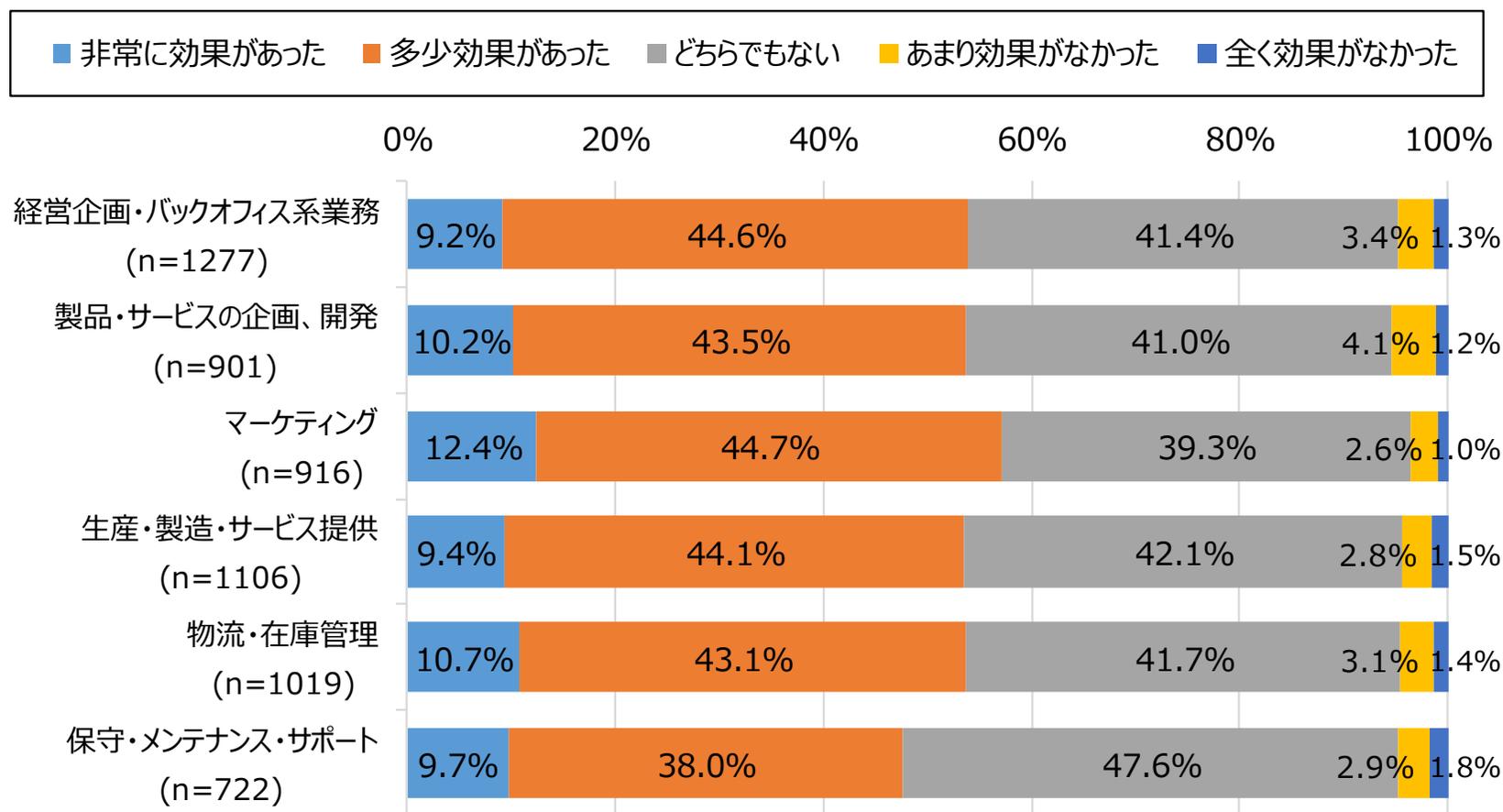
Q12：貴社では、データ活用ができるどのような環境を構築していますか？

※Q1で1つ以上選択肢1を選択した企業のみ回答



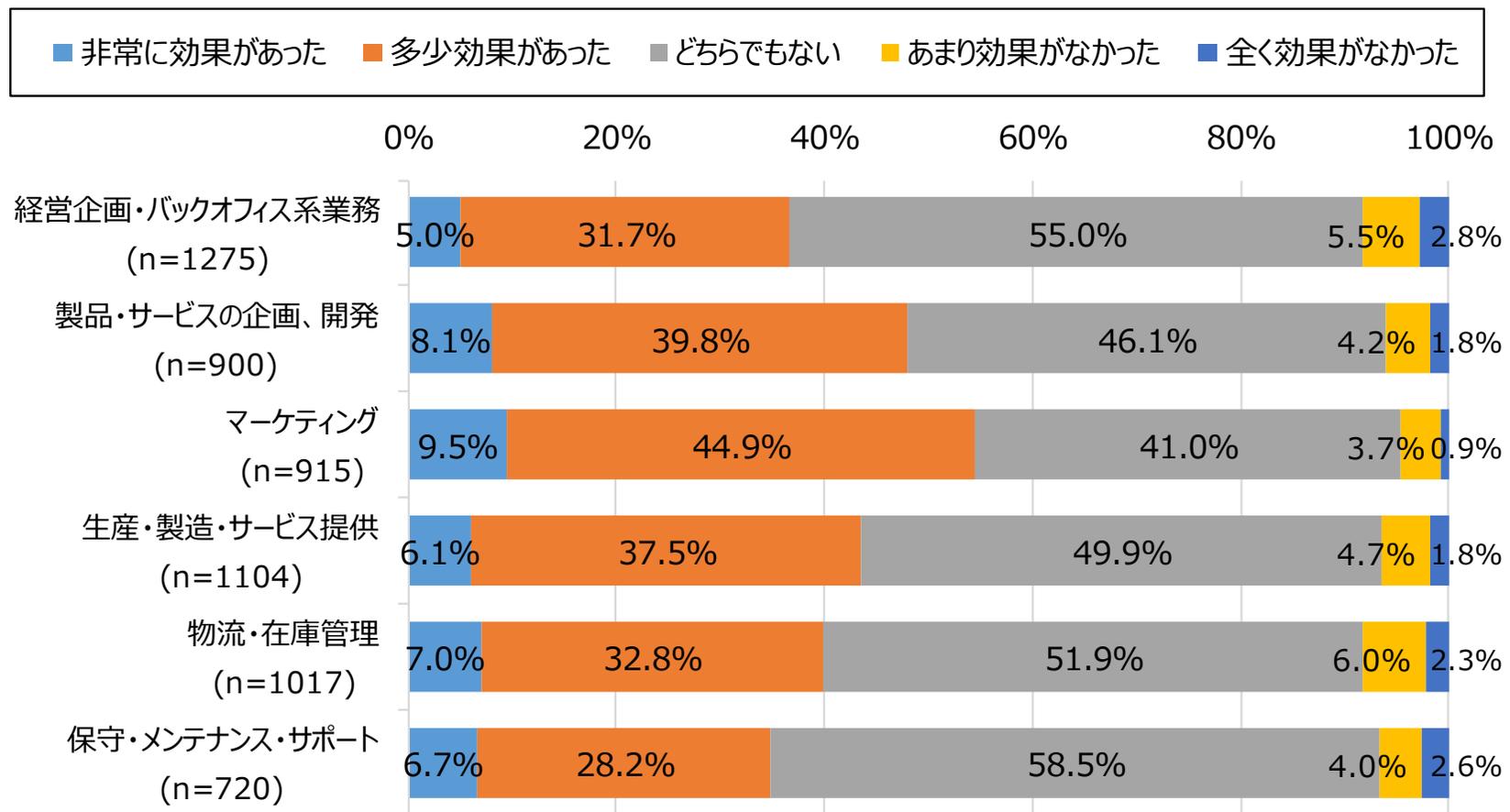
Q15：貴社では、それぞれの領域でデータを活用することによって、2019年度には投入面（業務効率化による費用削減等）でどのくらい効果がありましたか？

※Q1で選択肢1を選んだ項目のみを回答



Q16：貴社では、それぞれの領域でデータを活用することによって、2019年度には産出面（売上高増加等）でどのくらい効果がありましたか？

※Q1で選択肢1を選んだ項目のみを回答



企業が付加価値を生み出す生産要素として「資本」、「労働」及び「データ」を位置づけ、生産関数を推定。

(推定式)

$$V = A_o K^\alpha L^\beta Data^\gamma e^{dmy}$$

(変形後)

$$\longrightarrow \ln(V) = \ln A_o + \alpha \ln(K) + \beta \ln(L) + \gamma \ln(Data) + dmy$$

$$V = A_o K^\alpha L^\beta Data^\gamma e^{dmy}, \quad \alpha + \beta = 1 \quad \longrightarrow \quad \ln\left(\frac{V}{L}\right) = \ln A_o + \alpha \ln\left(\frac{K}{L}\right) + \gamma \ln(Data) + dmy$$

変数	計算方法等
付加価値 (V)	売上高 - 売上原価 - 販売費 - 一般管理費 + 減価償却費 ※企業活動基本調査より入手
資本 (K)	有形固定資産 + 無形固定資産 ※企業活動基本調査より入手
労働 (L)	常時従業員数 ※企業活動基本調査より入手
データ変数 (Data)	アンケート調査より入手

※ダミー変数としてindmy(製造業ダミー), sizedmy(大企業ダミー)を入れて推定した。

Data : 活用データ容量

$$\ln(V) = 0.461 + 0.39 \ln(K) + 0.503 \ln(L) + 0.035 \ln(Data) + 0.195 \text{indmy} - 0.196 \text{sizedmy}$$

(2.32) (18.76) (11.31) (2.71) (3.35) (-2.25)

adjR²=0.5330, () 内はt値, n=1417

$$\ln\left(\frac{V}{L}\right) = -0.06 + 0.39 \ln\left(\frac{K}{L}\right) + 0.026 \ln(Data) + 0.188 \text{indmy} - 0.378 \text{sizedmy}$$

(-1.09) (18.72) (2.12) (3.23) (-6.81)

adjR²=0.2554, () 内はt値, n=1417

ここで、Vは付加価値、Kは有形固定資産+無形固定資産、Lは常時従業員数、Dataは活用データ容量、indmyは製造業ダミー、sizedmyは大企業ダミーを表す。

Data : 内部入手した活用データ容量

$$\ln(V) = 0.461 + 0.39 \ln(K) + 0.504 \ln(L) + 0.032 \ln(Data) + 0.195 \text{indmy} - 0.196 \text{sizedmy}$$

(2.31)
(18.76)
(11.33)
(2.52)
(3.36)
(-2.26)

adjR²=0.5327, () 内はt値, n=1417

$$\ln\left(\frac{V}{L}\right) = -0.05 + 0.39 \ln\left(\frac{K}{L}\right) + 0.024 \ln(Data) + 0.188 \text{indmy} - 0.376 \text{sizedmy}$$

(-0.98)
(18.72)
(1.93)
(3.23)
(-6.77)

adjR²=0.2550, () 内はt値, n=1417

ここで、Vは付加価値、Kは有形固定資産+無形固定資産、Lは常時従業員数、Dataは内部入手した活用データ容量、indmyは製造業ダミー、sizedmyは大企業ダミーを表す。

Data : 1+外部入手した活用データ容量

$$\ln(V) = 0.412 + 0.39 \ln(K) + 0.515 \ln(L) + 0.083 \ln(Data) + 0.197indmy - 0.202sizedmy$$

(2.10) (18.76) (11.79) (3.10) (3.39) (-2.33)

adjR²=0.5338, () 内はt値, n=1417

$$\ln\left(\frac{V}{L}\right) = -0.06 + 0.39 \ln\left(\frac{K}{L}\right) + 0.075 \ln(Data) + 0.189indmy - 0.370sizedmy$$

(-1.08) (18.74) (2.82) (3.25) (-6.73)

adjR²=0.2573, () 内はt値, n=1417

※外部入手した活用データ容量は0の企業も多く分析対象が限られるため、1+外部入手活用データ容量を変数とした。ただし、活用データ容量が0の企業は分析対象から除いた。

ここで、Vは付加価値、Kは有形固定資産+無形固定資産、Lは常時従業員数、Dataは外部入手した活用データ容量、indmyは製造業ダミー、sizedmyは大企業ダミーを表す。

「データ活用の目的（Q20）」として多く挙げられている『業務の効率化』、『経営判断』、『マーケティング』において、それらで効果があった企業は、付加価値にもプラスの影響を与えているか。また、データ活用が効果を得るためには何が有効か。特に、「データ活用の課題（Q21）」として多く挙げられている『人材、スキル・ノウハウ不足』は重要な要素か。

①-1 データ活用の目的として多く挙がっている「経営企画・バックオフィス系業務」、「マーケティング」領域でデータを活用することによる主観的効果と付加価値との関係を確認
⇒ 正の関係性あり

①-2 責任者、データ活用環境、分析体制、分析人員（Q11～Q14）について、主観的効果・付加価値との関係を分析
⇒ <主観的効果> 分析では下記がプラスに有意

- ・ 「兼任の責任者」、「専任の責任者」
- ・ 「全社的」または「複数の部署内」でデータ活用ができる環境を構築
- ・ 「データ分析を行う専門部署の担当者」、「アライアンスやコンソーシアムなど他社等を交えた共同分析」
- ・ 分析人員数

⇒ <付加価値> 分析では下記がプラスに有意

- ・ 「全社的にデータ活用ができる環境を構築」
- ・ 「データ分析を行う専門部署の担当者」、「アライアンスやコンソーシアムなど他社等を交えた共同分析」

①-3 人材・スキル・ノウハウ不足の観点で、主観的効果・付加価値との関係を分析
⇒ <主観的効果> 分析では、下記がマイナスに有意

- ・ データ分析の体制（Q13）で「各事業部門のデータ分析が専門ではない人」のみを選択

⇒ <付加価値> 分析では、マイナスに有意となる項目はなかった

①-4 企業に対して、主観的効果、付加価値に対して有意にプラスに寄与しているものを推奨するとともに、有意にマイナスに寄与しているものの解消を促す

- ・ データを全社的に（少なくとも部を跨いで）共有・利用することの重要性
- ・ データ分析人材（部署）を確保・育成するとともに、産出面の効果を得るためには外部との連携した取組も重要
- ・ 人材の確保、スキルの向上は重要な要素であり、事例等を元に具体的な取組を紹介 など

主観的効果と各変数との関係性を順序プロビットモデルで分析。

$$S = \beta_0 + \sum_{i=1}^n \beta_i D_i + \varepsilon$$

ただし、 S 主観的効果、 D_i は企業*i*のデータ活用に関連する指標

被説明変数	計算方法
主観的効果（投入面）	選択肢1（非常に効果があった）を5点、選択肢2（多少効果があった）を4点、選択肢3（どちらでもない）を3点、選択肢4（あまり効果がなかった）を2点、選択肢5（全く効果がなかった）を1点
主観的効果（産出面）	
説明変数	計算方法 ※各項目0～1となるように基準化
責任者	Q11：選択肢をそれぞれダミー変数化
環境	Q12：選択肢をそれぞれダミー変数化
分析体制	Q13：選択肢をそれぞれダミー変数化
分析人員数	Q14：回答をそのまま用い、対数をとったものを変数化

主観的効果との関係を分析すると、「兼任の責任者」、「専任の責任者」がいること、「全社的」または「複数の部署内」でデータ活用ができる環境を構築していること、「データ分析を行う専門部署の担当者」による分析や「アライアンスやコンソーシアムなど他社等を交えた共同分析」、分析人員数がプラスに有意となった。

	経営企画・バックオフィス系業務
Q11_専任の責任者がいる	0.21*
Q11_兼任の責任者がいる	0.21***
Q12_全社的にデータ活用ができる環境を構築	0.37***
Q12_複数の部署内でデータ活用ができる環境を構築	0.22**
Q13_データ分析を行う専門部署の担当者	0.22***
Q13_各事業部門のデータ分析専門の担当者	0.11
Q13_アライアンスやコンソーシアムなど他社等を交えた共同分析	0.24
log(分析人員数)	0.06**
製造業ダミー	-0.06
大企業ダミー	0.02
修正R ²	0.0287
n数	990

	マーケティング
Q11_専任の責任者がいる	0.26*
Q11_兼任の責任者がいる	0.25***
Q12_全社的にデータ活用ができる環境を構築	0.30***
Q12_複数の部署内でデータ活用ができる環境を構築	0.24**
Q13_データ分析を行う専門部署の担当者	0.31***
Q13_各事業部門のデータ分析専門の担当者	0.07
Q13_アライアンスやコンソーシアムなど他社等を交えた共同分析	0.48**
log(分析人員数)	0.09**
製造業ダミー	-0.23**
大企業ダミー	0.09
修正R ²	0.0450
n数	724

付加価値との関係を分析すると、「全社的にデータ活用ができる環境を構築」していること、「データ分析を行う専門部署の担当者」による分析や「アライアンスやコンソーシアムなど他社等を交えた共同分析」がプラスに有意となった。

<一次同次なし>

	推定結果
log(有形固定資産 + 無形固定資産)	0.38***
log(常時従業員数)	0.48***
log(活用データ容量)	0.03**
Q11_専任の責任者がいる	0.003
Q11_兼任の責任者がいる	0.05
Q12_全社的にデータ活用ができる環境を構築	0.16*
Q12_複数の部署内でデータ活用ができる環境を構築	0.04
Q13_データ分析を行う専門部署の担当者	0.15*
Q13_各事業部門のデータ分析専門の担当者	0.04
Q13_アライアンスやコンソーシアムなど他社等を交えた共同分析	0.49**
log(分析人員数)	0.04
製造業ダミー	0.21***
大企業ダミー	-0.21**
修正R ²	0.5387
n数	1019

<一次同次あり>

	推定結果
log(有形固定資産 + 無形固定資産)	0.38***
log(常時従業員数)	(0.62) ※逆算
log(活用データ容量)	0.04**
Q11_専任の責任者がいる	-0.02
Q11_兼任の責任者がいる	0.04
Q12_全社的にデータ活用ができる環境を構築	0.17**
Q12_複数の部署内でデータ活用ができる環境を構築	0.03
Q13_データ分析を行う専門部署の担当者	0.13*
Q13_各事業部門のデータ分析専門の担当者	0.03
Q13_アライアンスやコンソーシアムなど他社等を交えた共同分析	0.45**
log(分析人員数)	0.03
製造業ダミー	0.20***
大企業ダミー	-0.38***
修正R ²	0.2603
n数	1019

社内データだけではなく、外部から入手したデータを活用することの有用性を示すことができるか。

②-1 まずは、外部入手データの活用状況について確認。

- ⇒ データを活用している企業の内、約半数の企業で外部入手データを活用している。
- ⇒ 有償データに限定した場合、特にサービス業で活用割合が低下する。

②-2 外部入手データの活用状況（Q8）と主観的効果・付加価値との関係を分析

- ⇒ <主観的効果> 分析では「外部データの利用あり（有償、無償問わず）」が有意にプラス
- ⇒ <付加価値> 分析では「外部データの利用あり（有償）」が有意にプラス

②-3 活用データ容量、内部入手活用データ容量、外部入手活用データ容量を加えた生産関数モデルを推定。

- ⇒ 活用データ容量、内部入手活用データ容量、外部入手活用データ容量はいずれも生産要素としてプラスに貢献し、外部入手活用データ容量の係数が最も大きい。

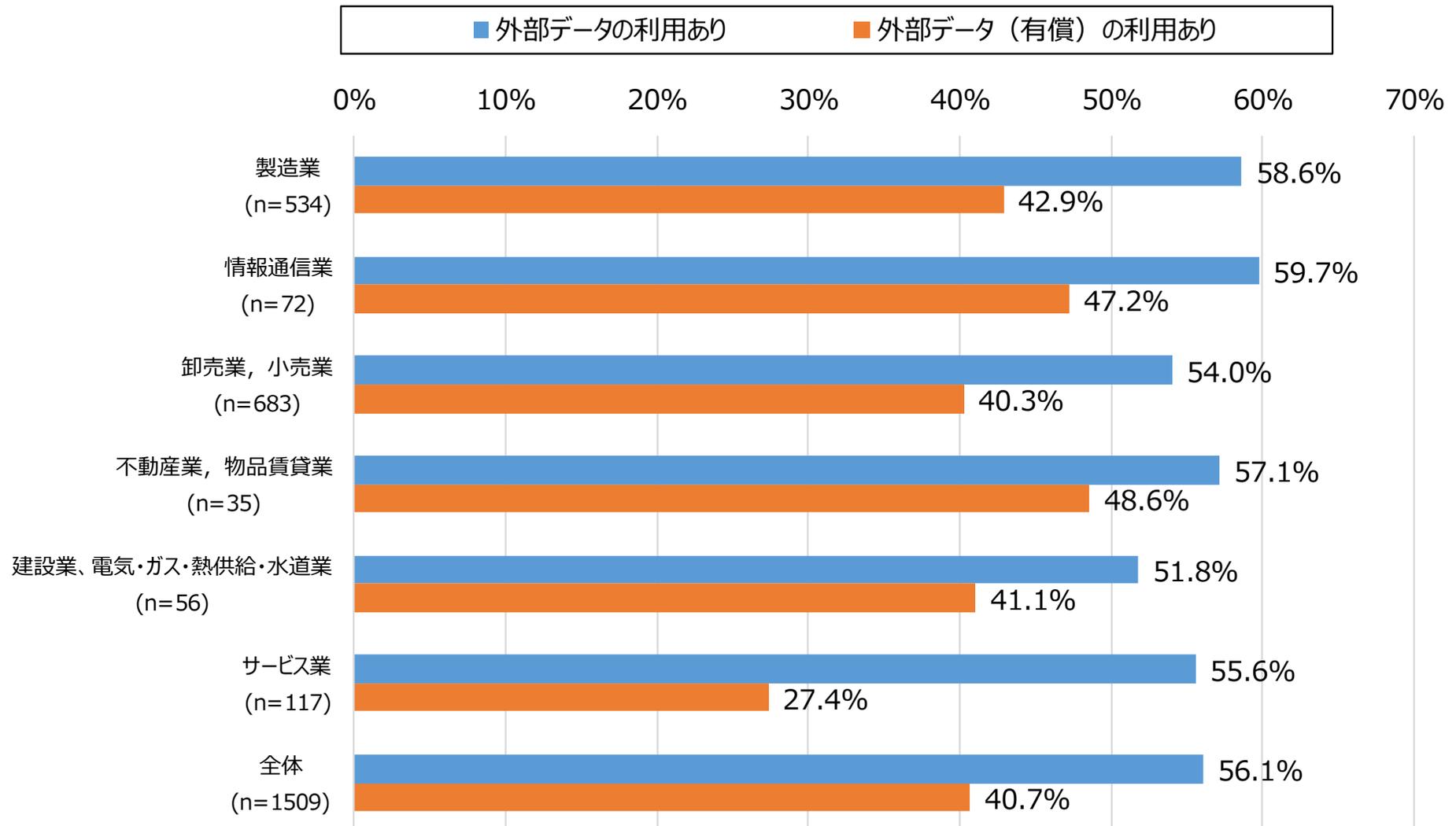
分析結果から、外部入手データは内部入手データとは別にプラスの貢献をしていると考えられ、外部から入手するデータによって社内システムやデータが有機的に結びつきデータ活用の効果が加速すると言える。

以上より、外部データを入手することを容易にする取組（データ取引市場等）を促進

- ⇒ どこにどのようなデータがあるのか、安心・安全なデータ取引の確立、値付け方法 等

Q8：貴社で活用しているデータのうち、1年間（2019年度）に新たに外部から提供されたものはありませんか。あれば入手元の種類、国内・国外、有償・無償別に、当てはまるものをすべてお答えください。

※Q1で1つ以上選択肢1を選択した企業のみ回答（n=1,509）



付加価値との関係を分析すると、「有償外部データの利用あり」がプラスに有意となった。

<一次同次なし>

	推定結果
log(有形固定資産 + 無形固定資産)	0.37***
log(常時従業員数)	0.52***
log(活用データ容量)	0.02
Q8_有償外部データの利用あり	0.25***
Q8_無償外部データの利用あり	0.03
Q11_専任の責任者がいる	-0.01
Q11_兼任の責任者がいる	0.07
Q12_全社的にデータ活用ができる環境を構築	0.12*
Q12_複数の部署内でデータ活用ができる環境を構築	-0.03
Q13_データ分析を行う専門部署の担当者	0.14*
Q13_各事業部門のデータ分析専門の担当者	0.04
Q13_アライアンスやコンソーシアムなど他社等を交えた共同分析	0.33*
製造業ダミー	0.23***
大企業ダミー	-0.22**
修正R ²	0.5399
n数	1300

<一次同次あり>

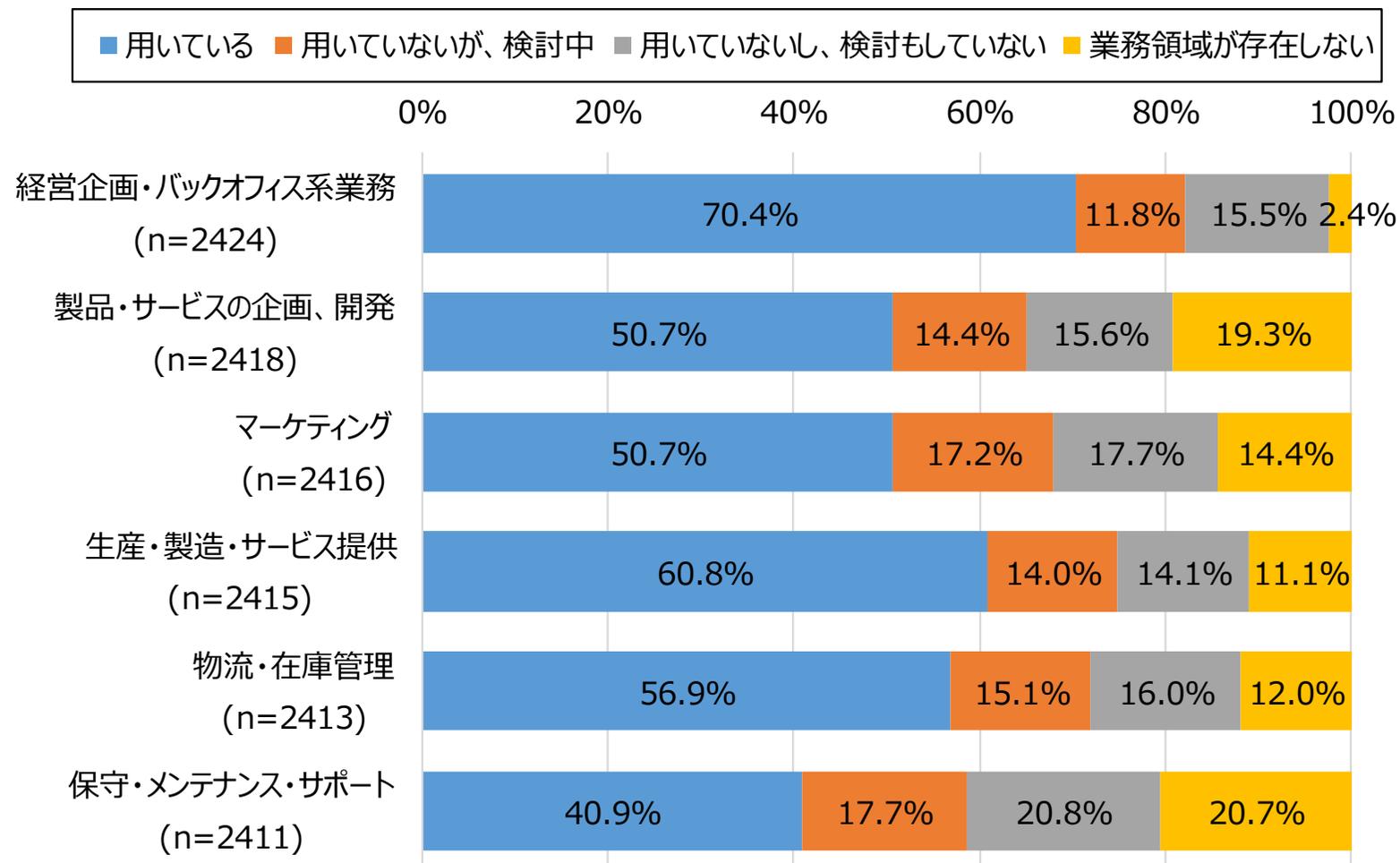
	推定結果
log(有形固定資産 + 無形固定資産)	0.37***
log(常時従業員数)	(0.63) ※逆算
log(活用データ容量)	0.02
Q8_有償外部データの利用あり	0.25***
Q8_無償外部データの利用あり	0.02
Q11_専任の責任者がいる	-0.03
Q11_兼任の責任者がいる	0.06
Q12_全社的にデータ活用ができる環境を構築	0.13*
Q12_複数の部署内でデータ活用ができる環境を構築	-0.03
Q13_データ分析を行う専門部署の担当者	0.13*
Q13_各事業部門のデータ分析専門の担当者	0.02
Q13_アライアンスやコンソーシアムなど他社等を交えた共同分析	0.29
製造業ダミー	0.22***
大企業ダミー	-0.38***
修正R ²	0.2713
n数	1300

(注) *有意水準10%、**有意水準5%、***有意水準1%

(参考資料)

1. アンケート調査結果

Q1：貴社では、それぞれの業務領域において、データ分析した結果を何らかの判断に用いていますか？

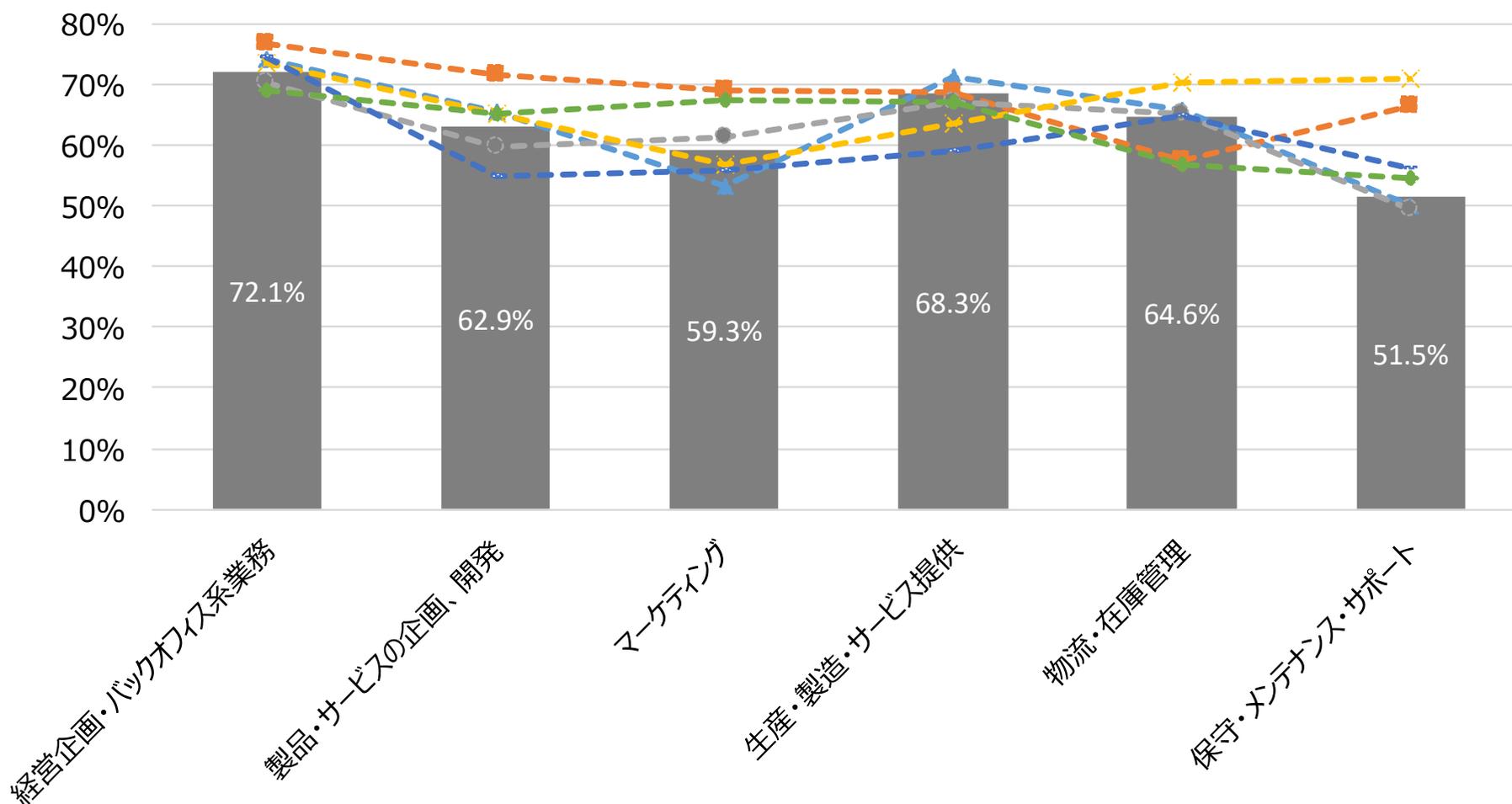
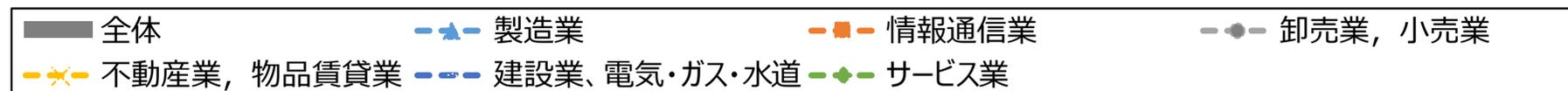


Q1：貴社では、それぞれの業務領域において、データ分析した結果を何らかの判断に用いていますか？

※業務領域が存在する企業のみで集計

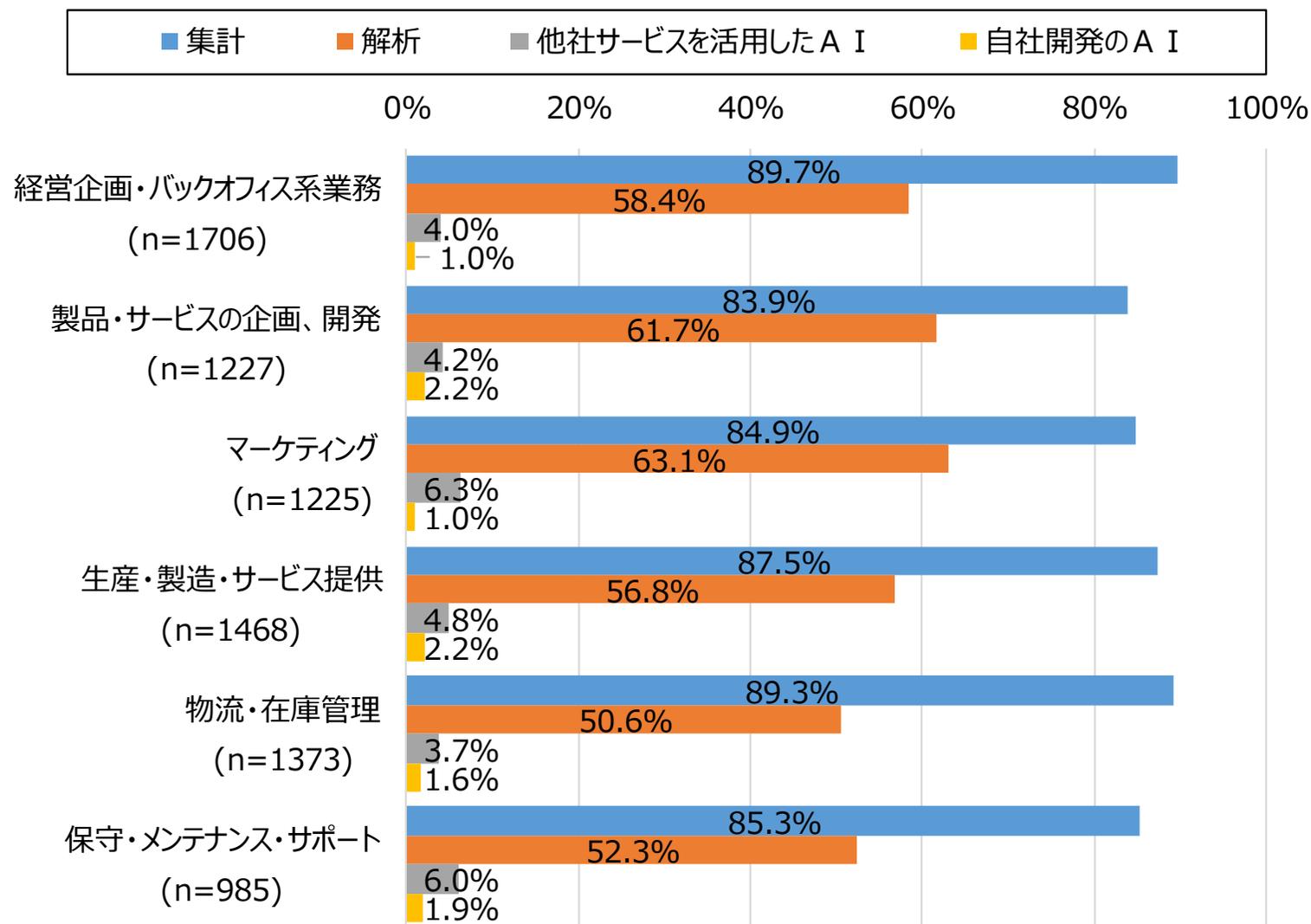
※「用いている」の割合

サービス業は、日本標準産業分類における「学术研究，専門・技術サービス業」、「宿泊業，飲食サービス業」、「生活関連サービス業，娯楽業」、「教育，学習支援業」、「医療，福祉」、「複合サービス事業」、「サービス業（他に分類されないもの）」



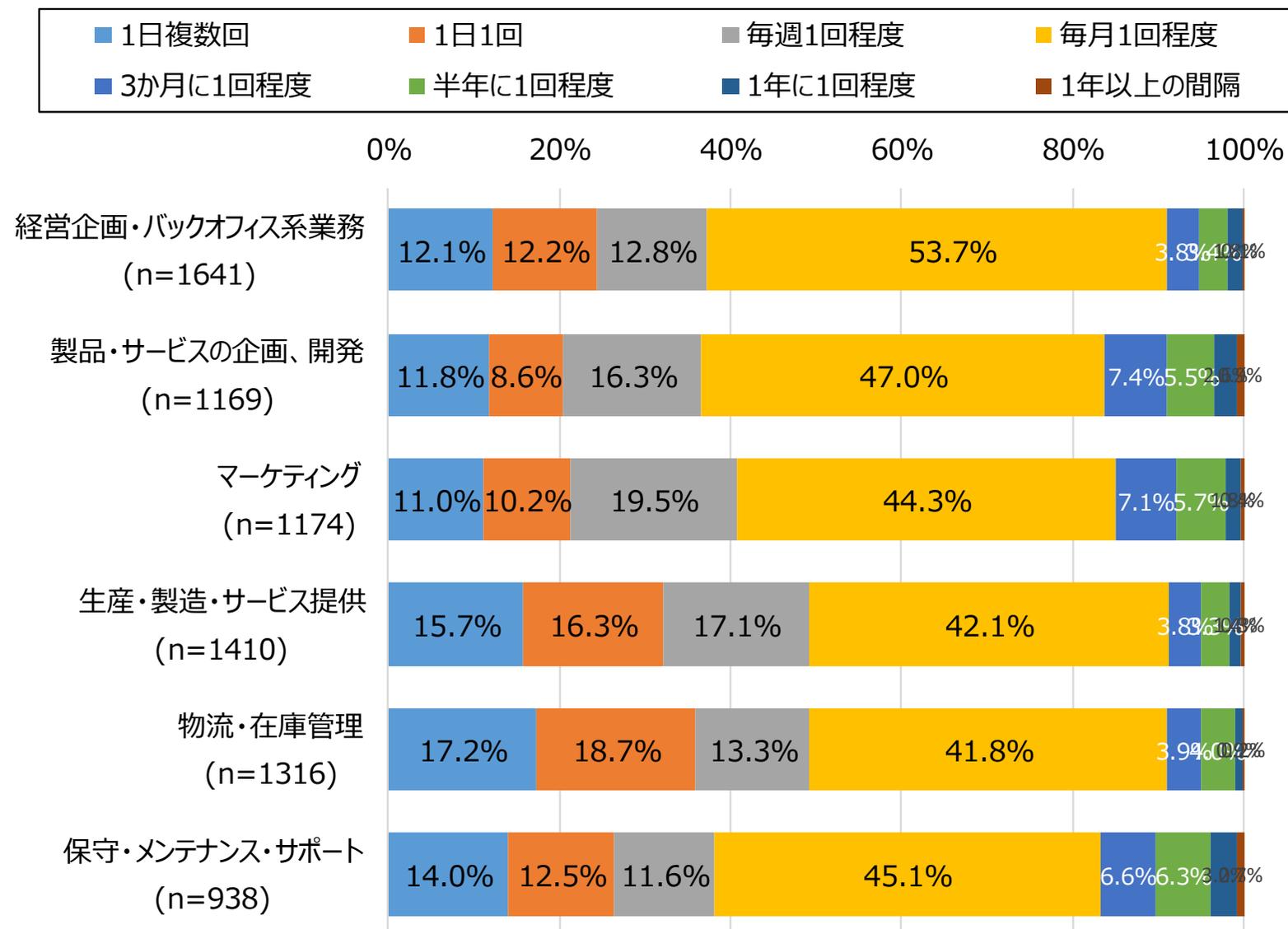
Q2：貴社では、データをどのように処理したものを各領域で活用していますか？当てはまるものすべてをお答えください。

※Q1で選択肢1を選んだ項目のみを回答



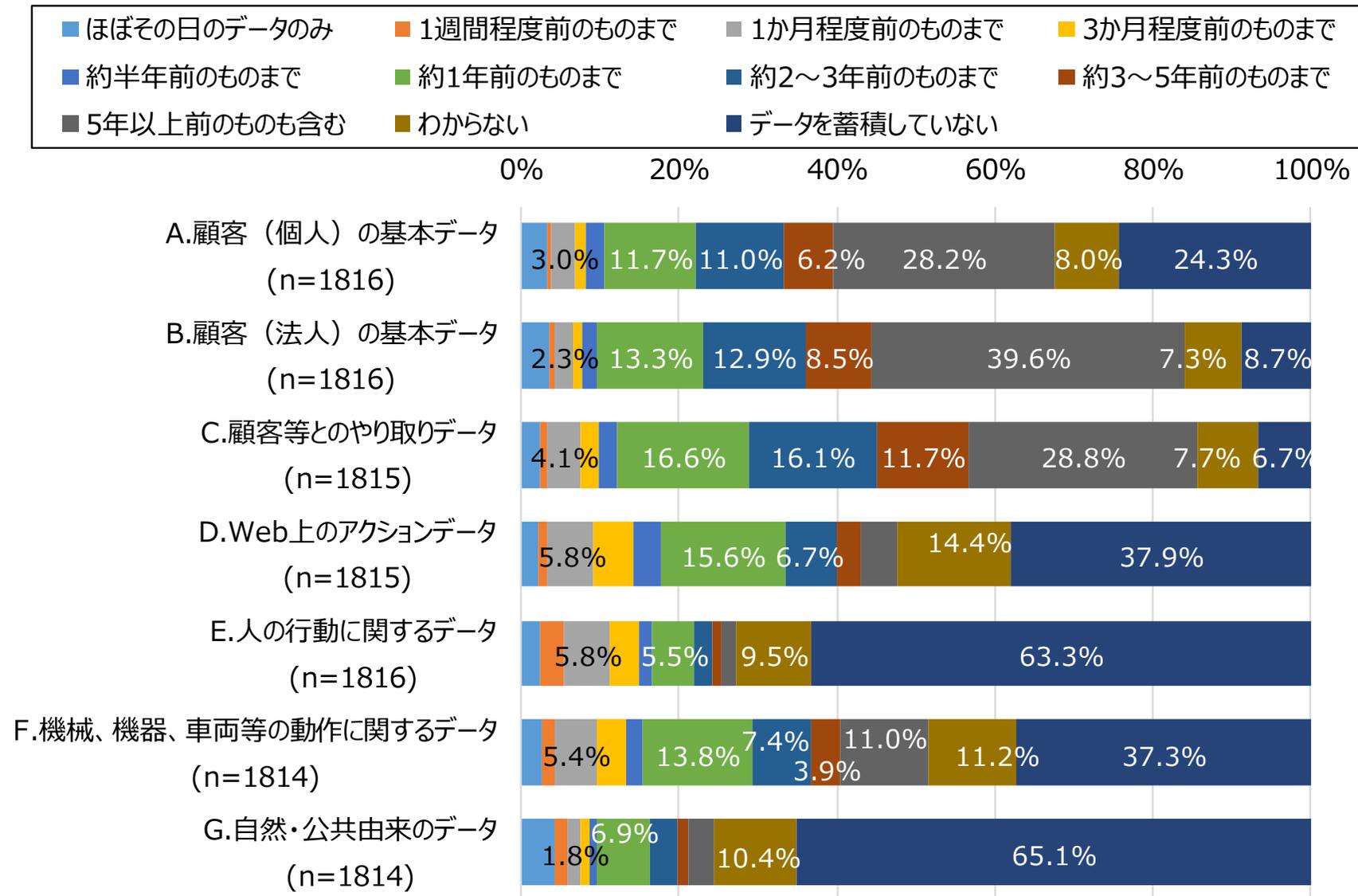
Q3：貴社では、データ分析をどのくらいの頻度（間隔）で行いますか？

※Q1で選択肢1を選んだ項目のみを回答



Q4：貴社では、データ分析を行う際、おおよそどれくらい前に蓄積したデータまで遡って利用しますか？

※Q1で1つ以上選択肢1を選択した企業のみ回答



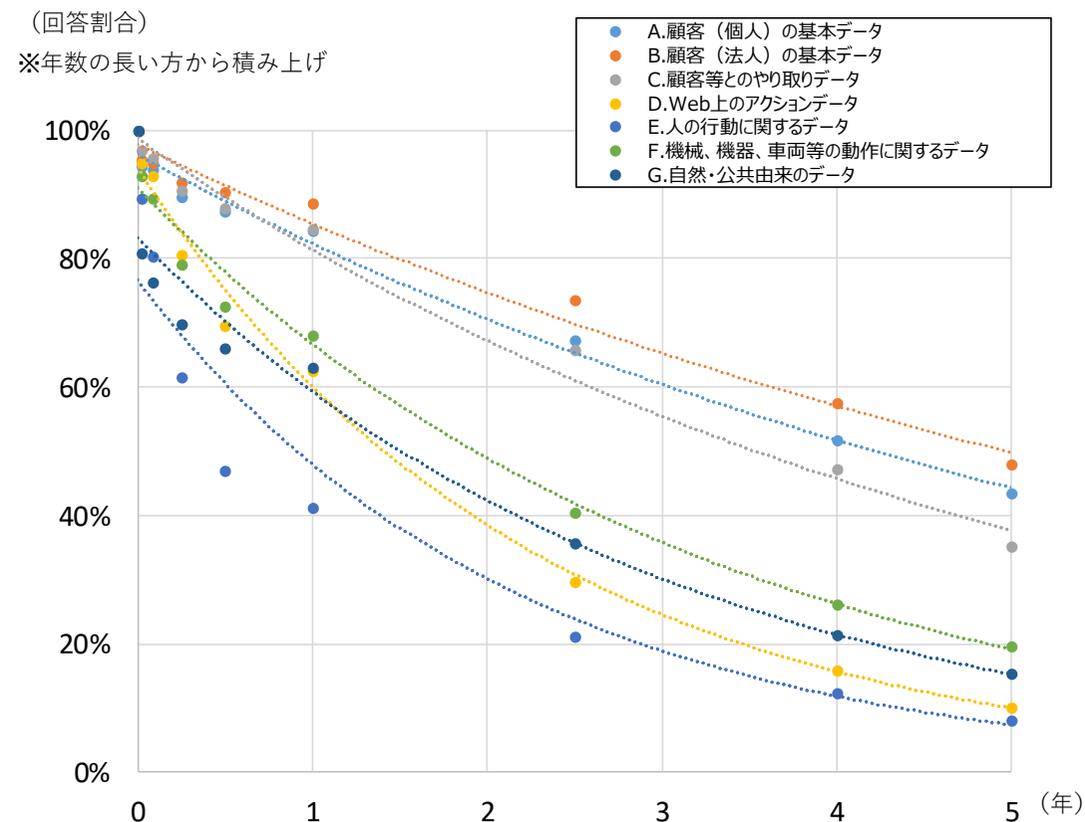
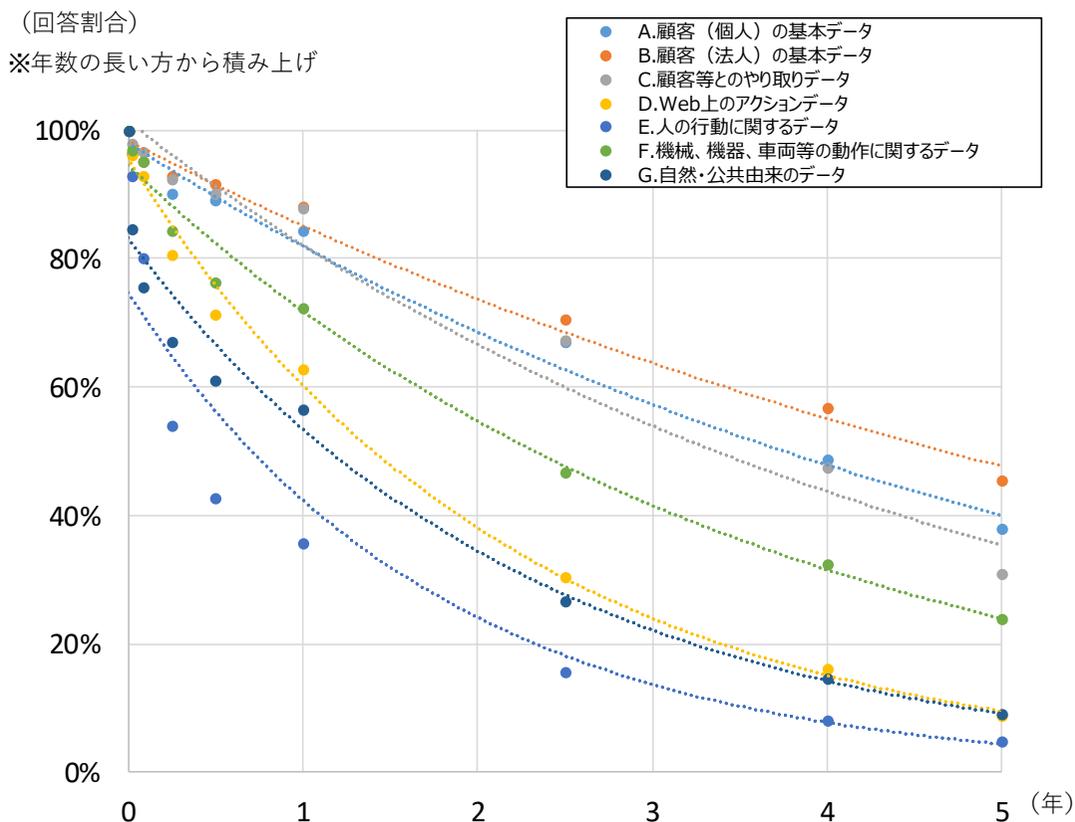
製造業、非製造業に分けて観察すると、

- 顧客関係のデータについては、製造業・非製造業で大きな差異はみられない。
- 「機械、機器、車両等の動作に関するデータ」は非製造業の方が減耗が大きい傾向がみられるが、極端に大きな差異ではない。

※「機械、機器、車両等の動作に関するデータ」については、製造業と非製造業で利用用途が異なることも考えられる。

<製造業>

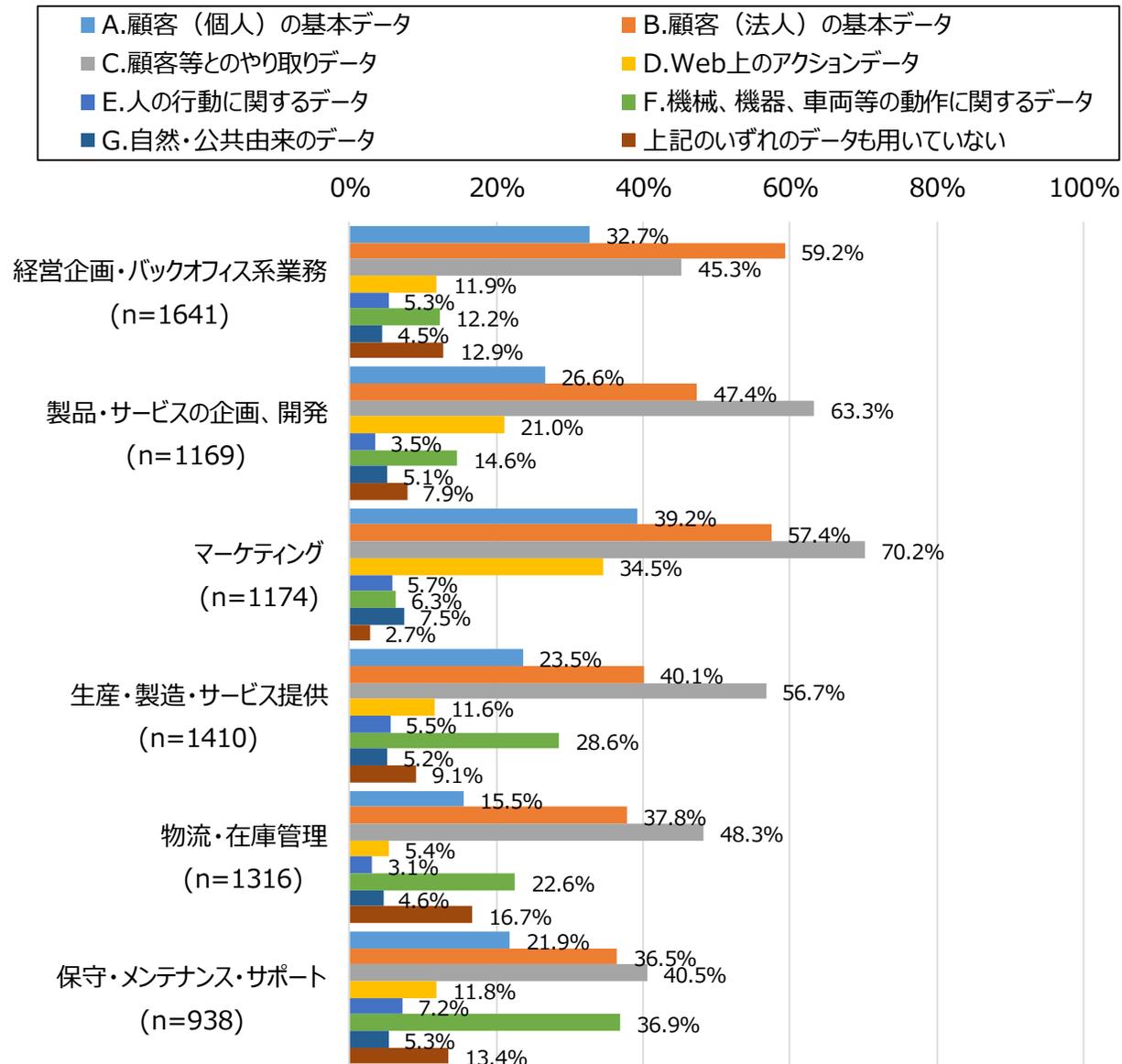
<非製造業>



※「わからない」を除いて集計。

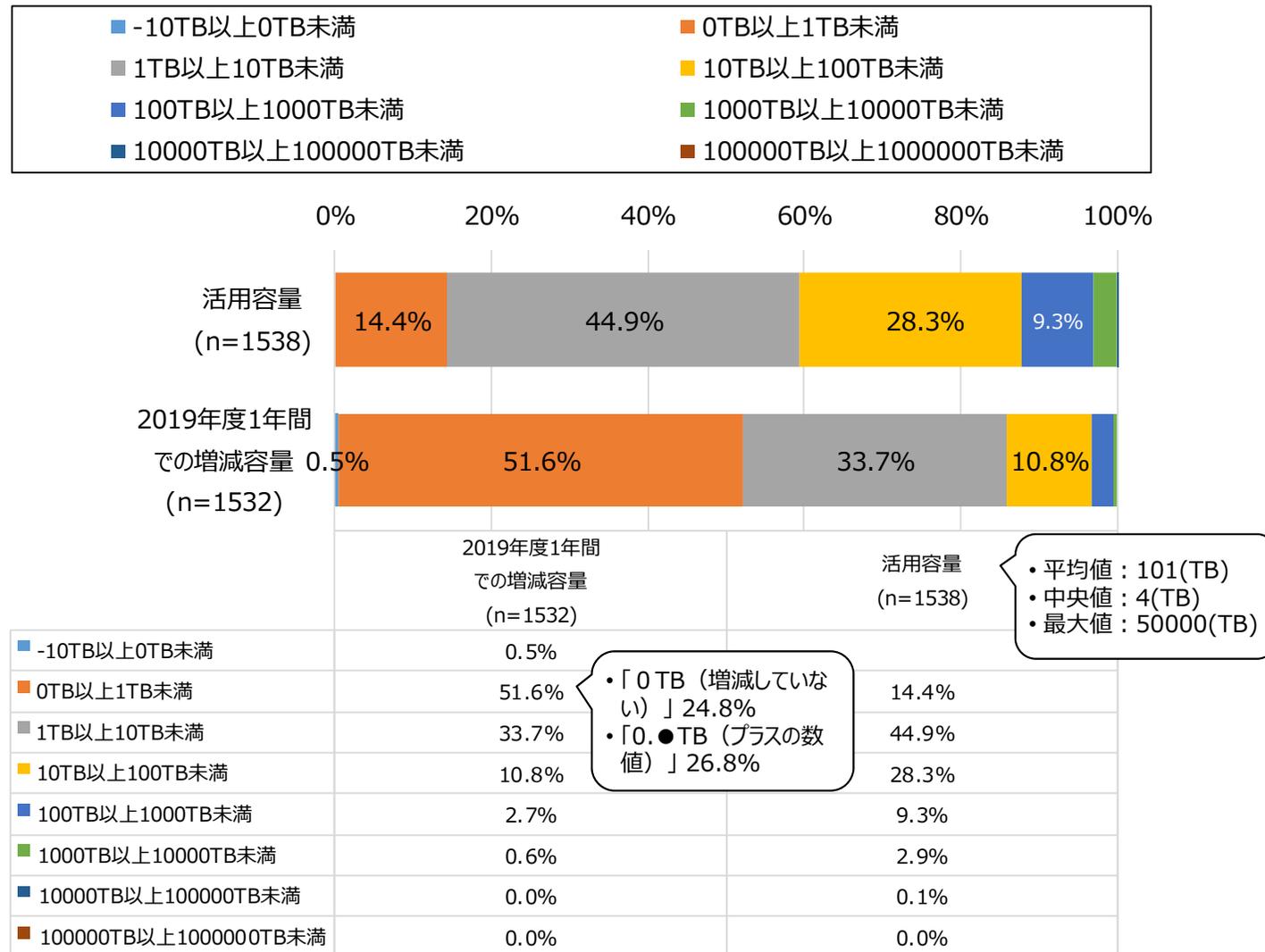
Q5：貴社では、それぞれの領域においてデータ分析をする際にどのようなデータを用いていますか？

※Q1で選択肢1を選んだ項目のみを回答



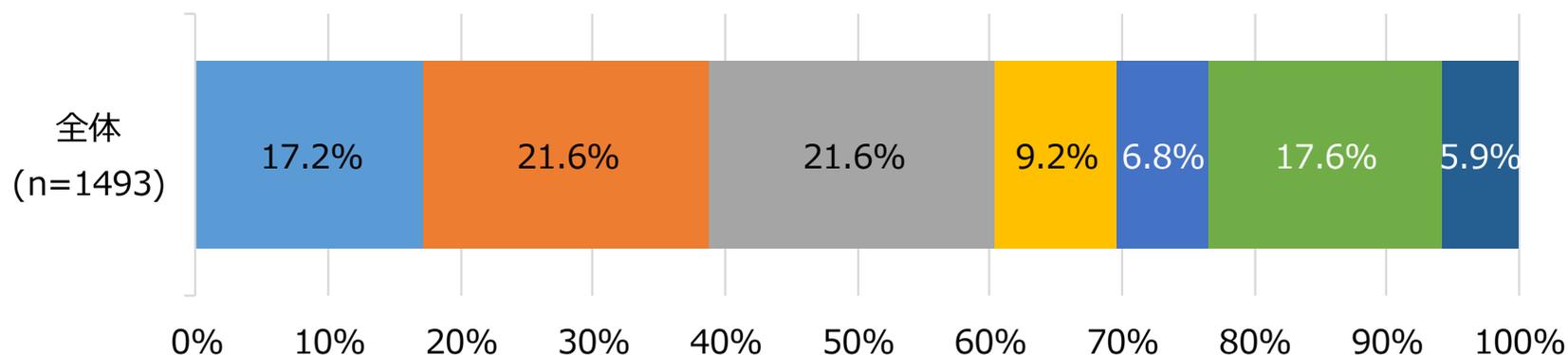
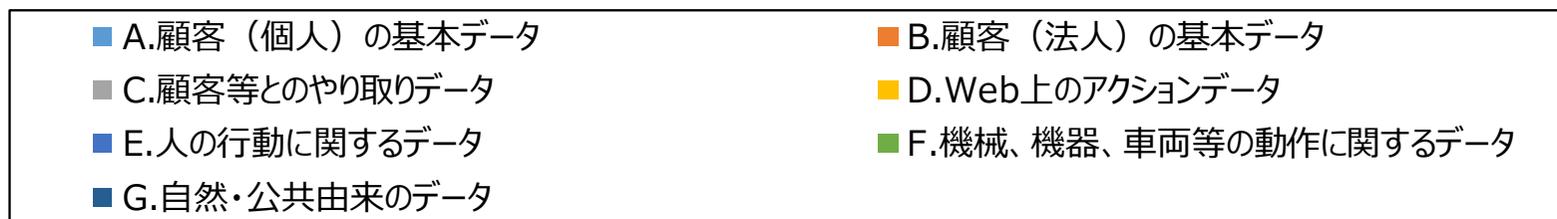
Q6：貴社で活用しているデータについて、2019年度末時点における容量の総計をお答えください。また、2019年度の1年間でどの程度増減したかお答えください（減少した場合はマイナスでご記入ください）。

※Q1で1つ以上選択肢1を選択した企業のみ回答



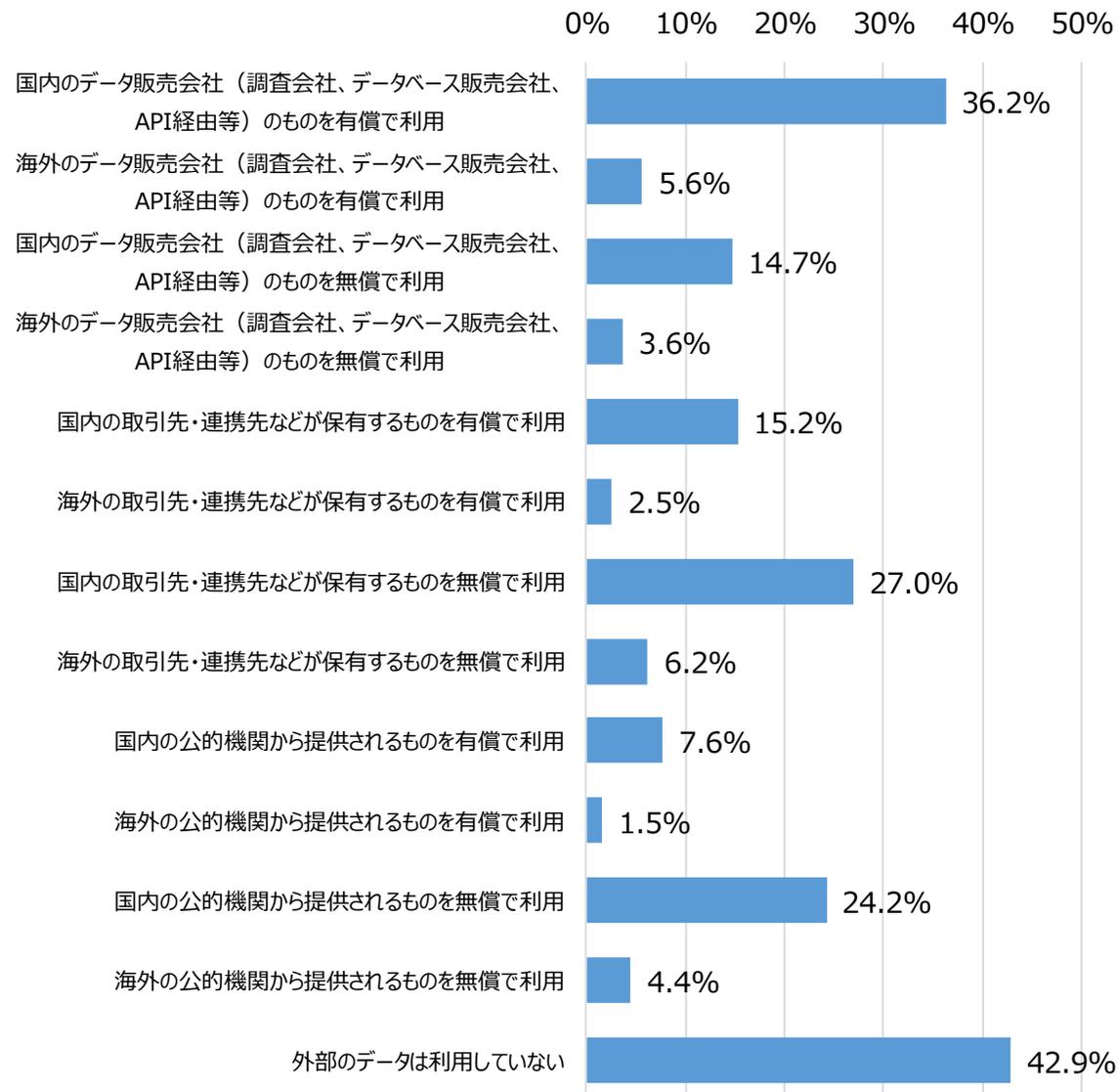
Q7：貴社で活用しているデータのうち、それぞれの容量割合はおよそどれくらいですか？合計が100%になるようにお答えください。

※Q1で1つ以上選択肢1を選択かつQ4いずれかの項目で選択肢1～10を選択した企業のみ回答



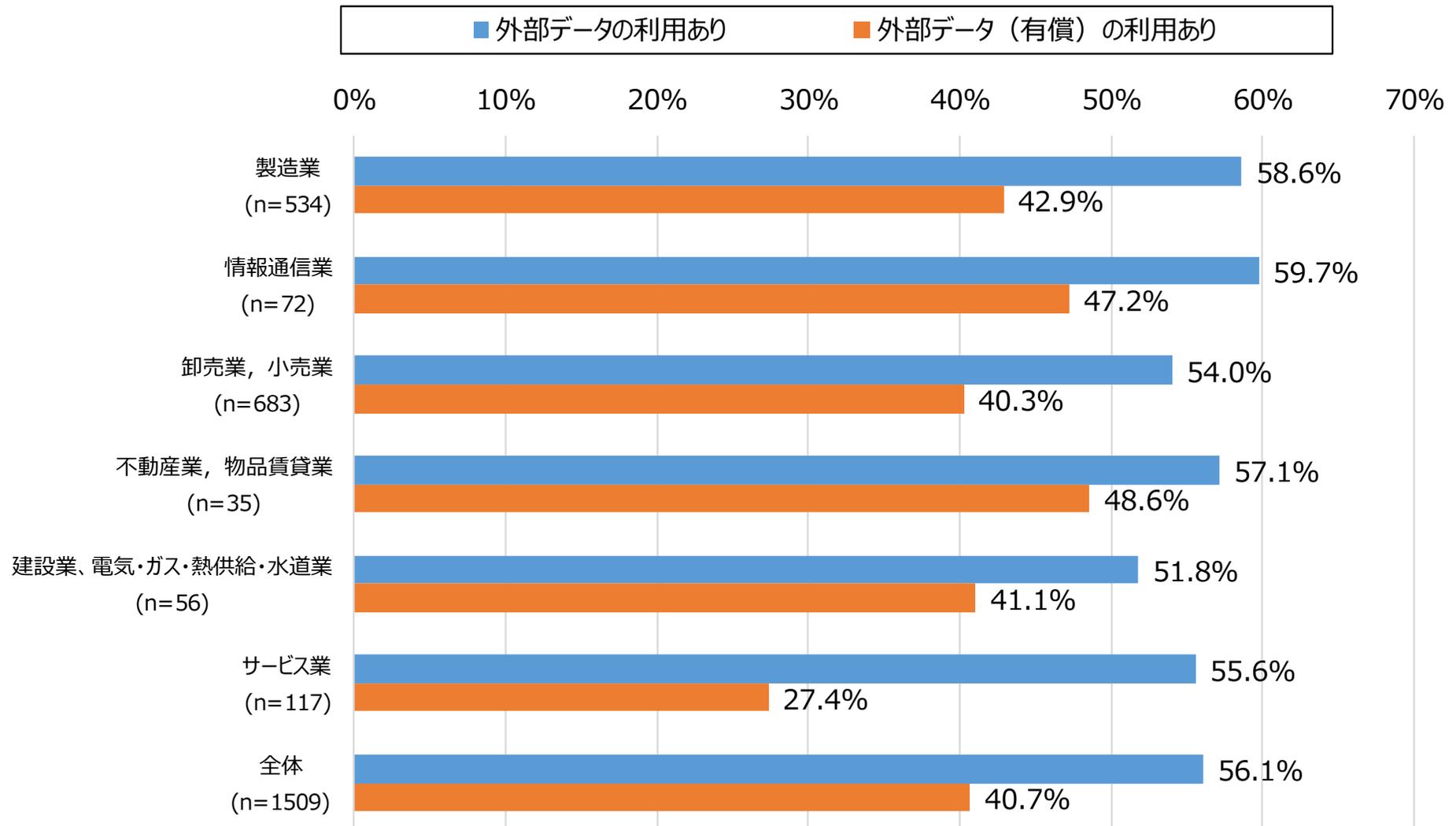
Q8：貴社で活用しているデータのうち、1年間（2019年度）に新たに外部から提供されたものはありませんか。あれば入手元の種類、国内・国外、有償・無償別に、当てはまるものをすべてお答えください。

※Q1で1つ以上選択肢1を選択した企業のみ回答（n=1,509）



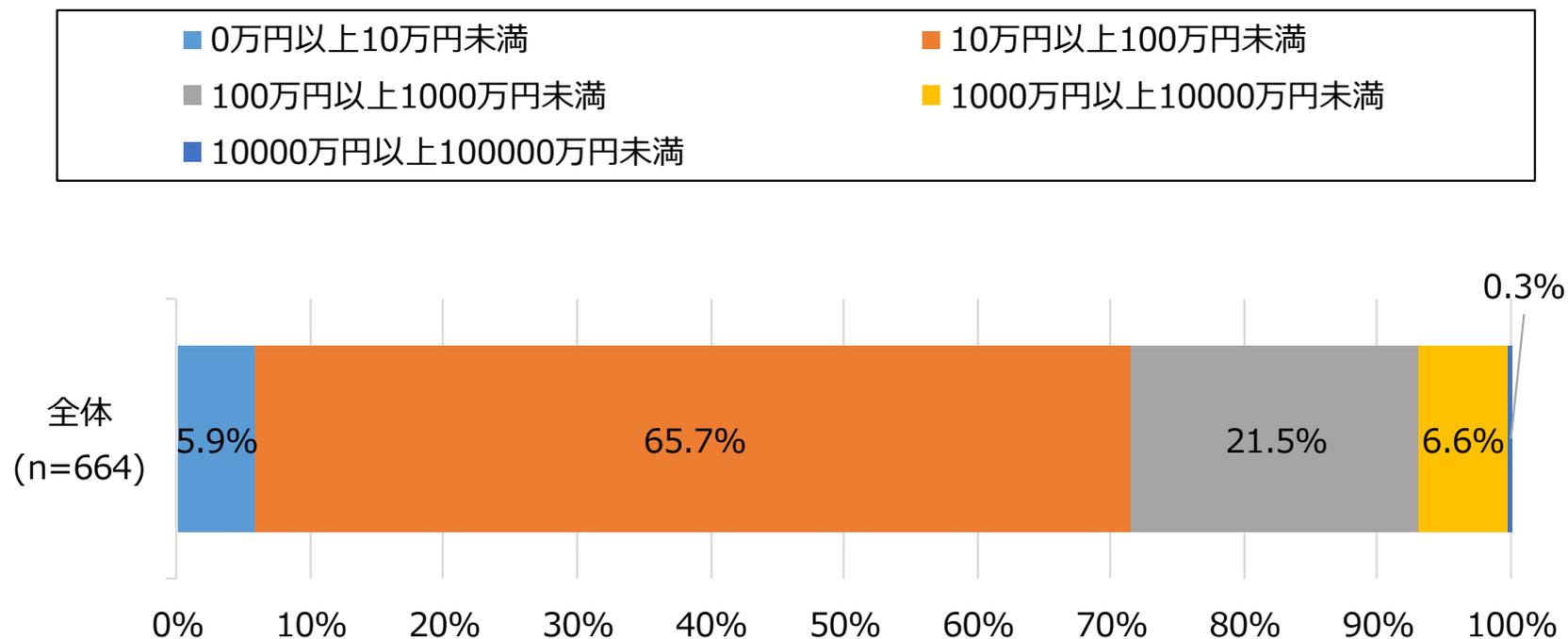
Q8：貴社で活用しているデータのうち、1年間（2019年度）に新たに外部から提供されたものがありますか。あれば入手元の種類、国内・国外、有償・無償別に、当てはまるものをすべてお答えください。

※Q1で1つ以上選択肢1を選択した企業のみ回答（n=1,509）

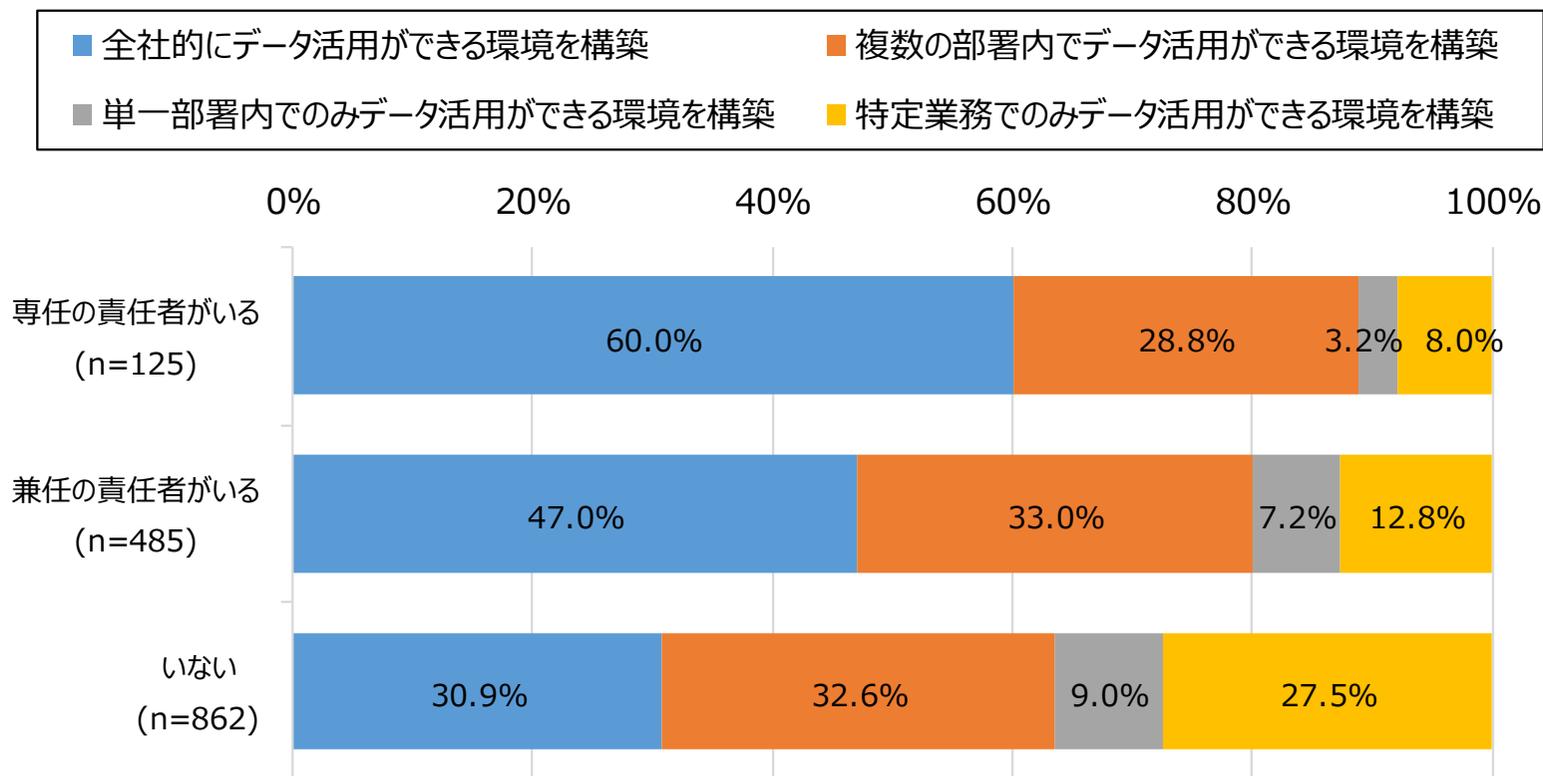


Q9：先の設問Q8で示した、外部から提供されたものに対して、貴社では1年間（2019年度）にどれだけ支出をしましたか。総額を万円の単位でお答えください。

※Q8で選択肢1,2,5,6,9,10のいずれかを選択した企業

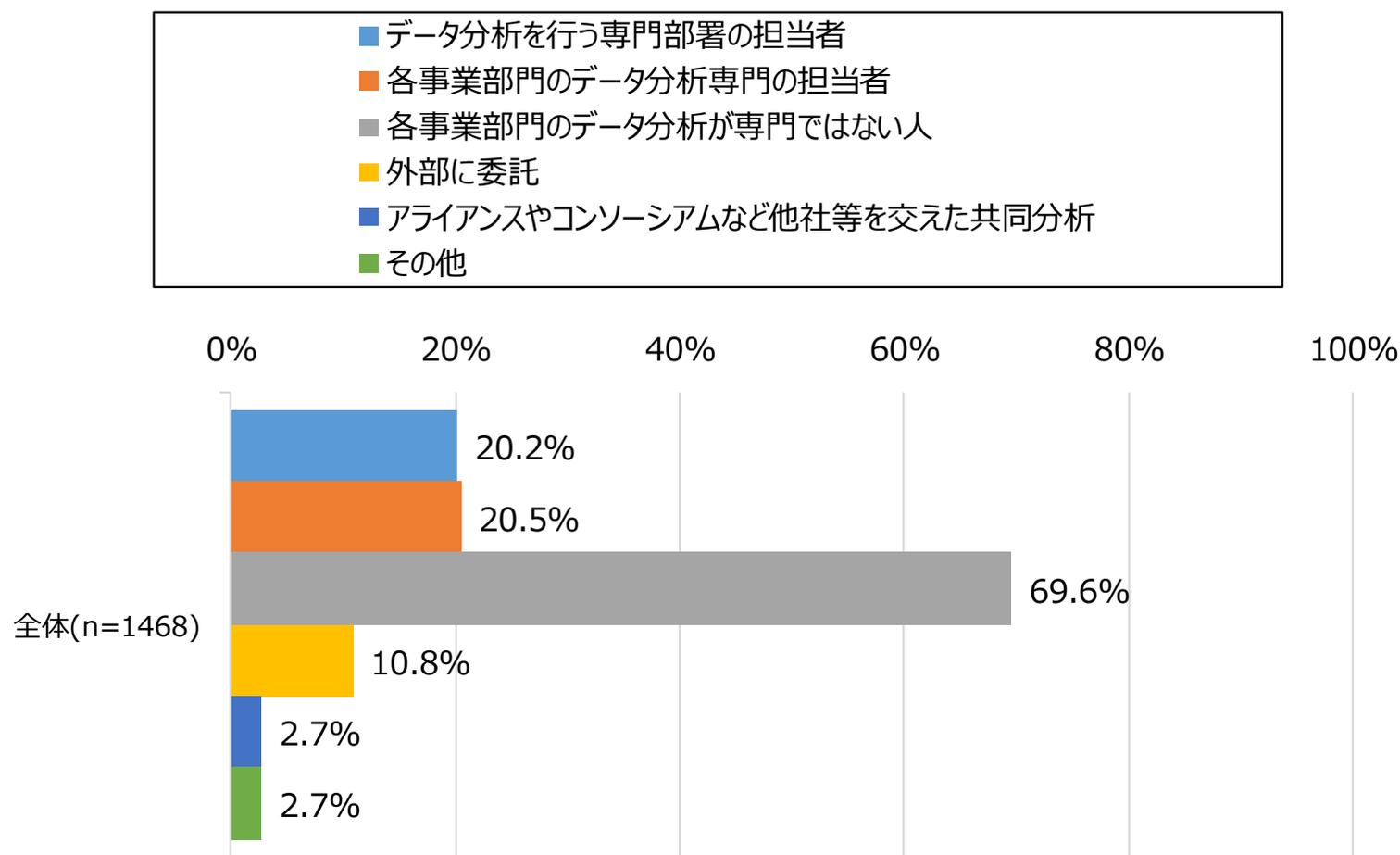


データ活用を主導する責任者と環境構築との関係を見ると、責任者がいる企業（さらに、兼任よりも専任）の方が「全社的にデータ活用ができる環境を構築」している企業の割合が大きい。



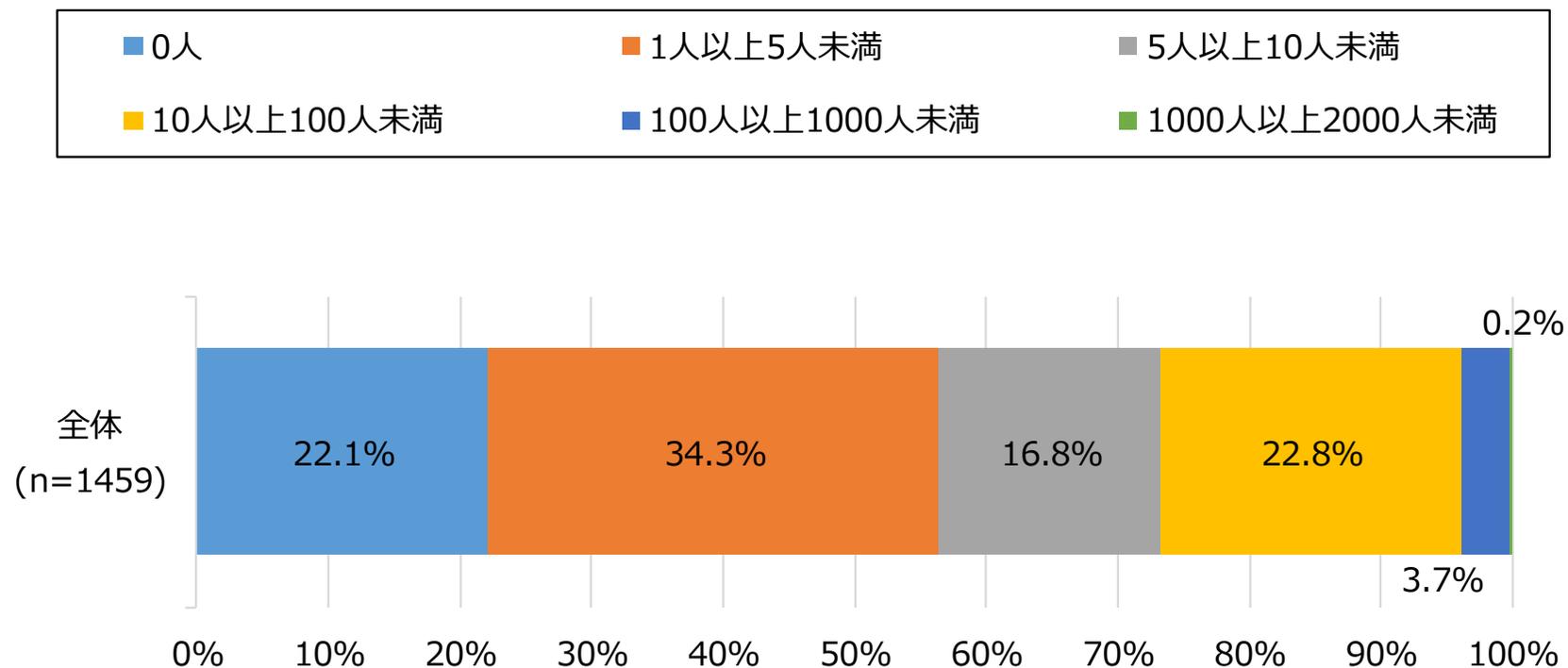
Q13：貴社がデータ分析を行う際、どのような体制（部署・担当者）で分析（外部企業等が提供しているデータ解析ツール・サービスの利用も含みます）を行いますか？

※Q1で1つ以上選択肢1を選択した企業のみ回答

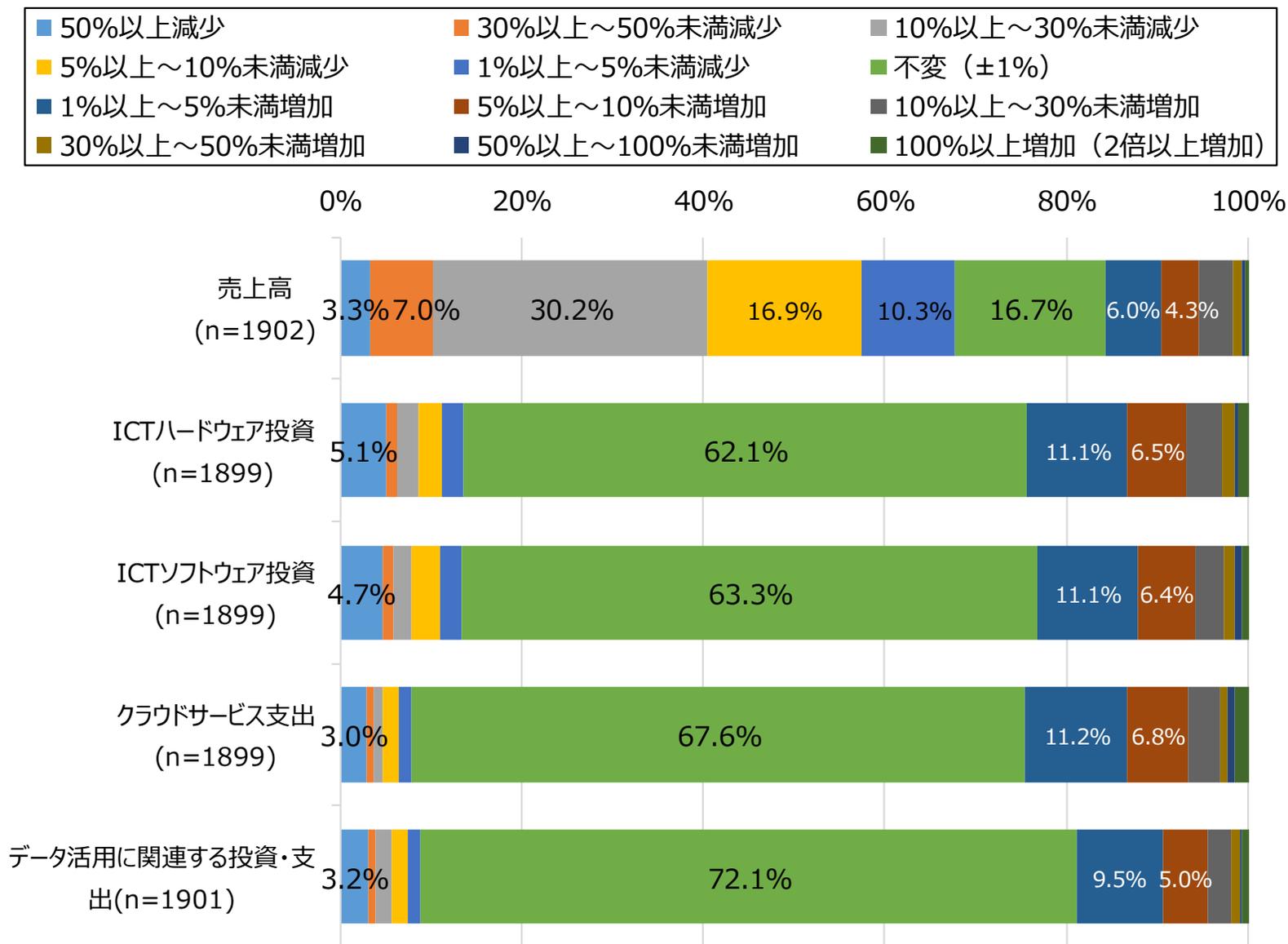


Q14：貴社で、データ分析を主な業務とする従業員（正社員だけではなく、パートタイマー・アルバイト・契約社員・派遣社員等を含みます）はおおよそ何人くらいいますか？

※Q1で1つ以上選択肢1を選択した企業のみ回答

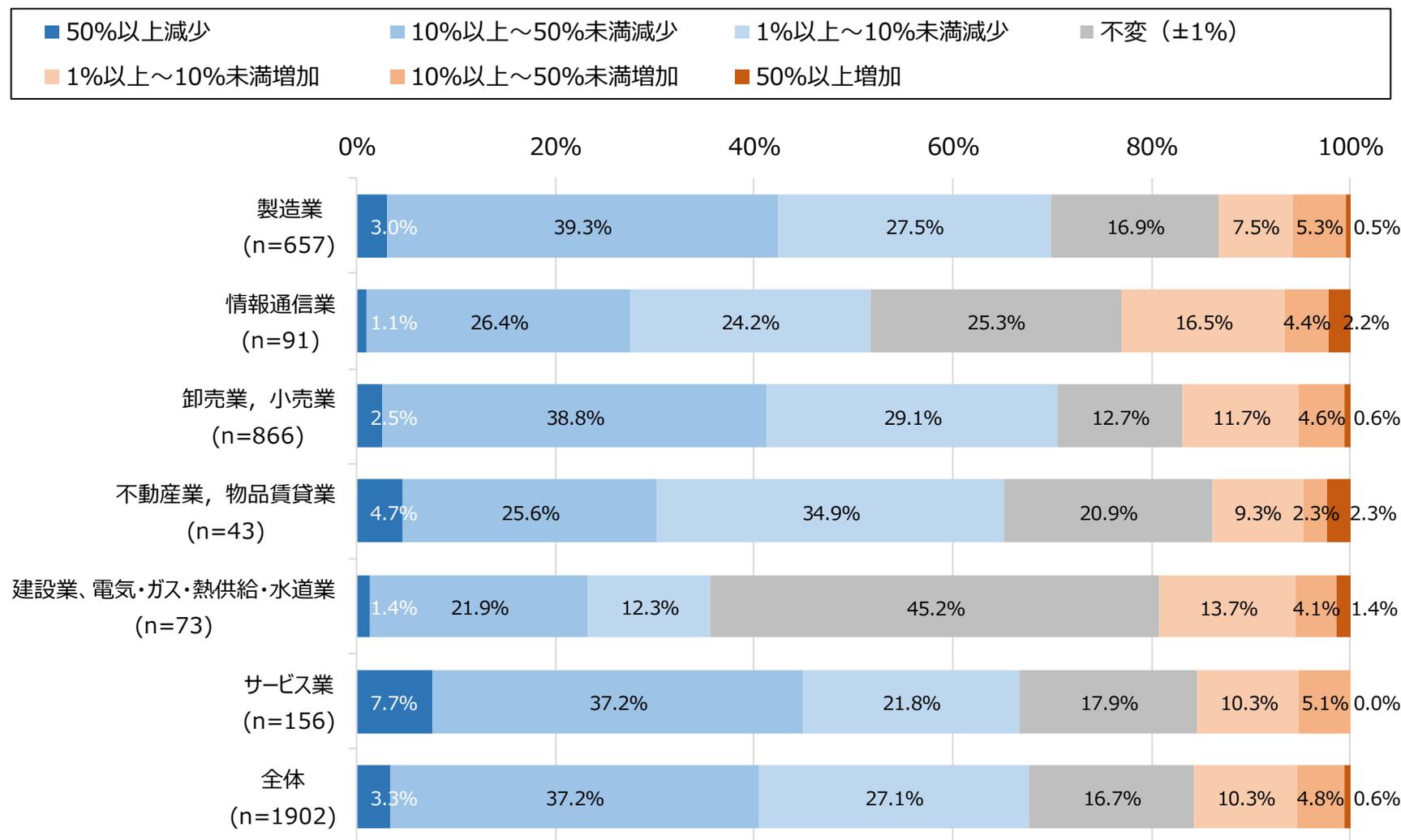


Q17：新型コロナウイルス感染症の影響により、貴社の経営等にどのような影響（変化）がありましたか？2019年度と比べて2020年度がどのように変化したかをお答えください。



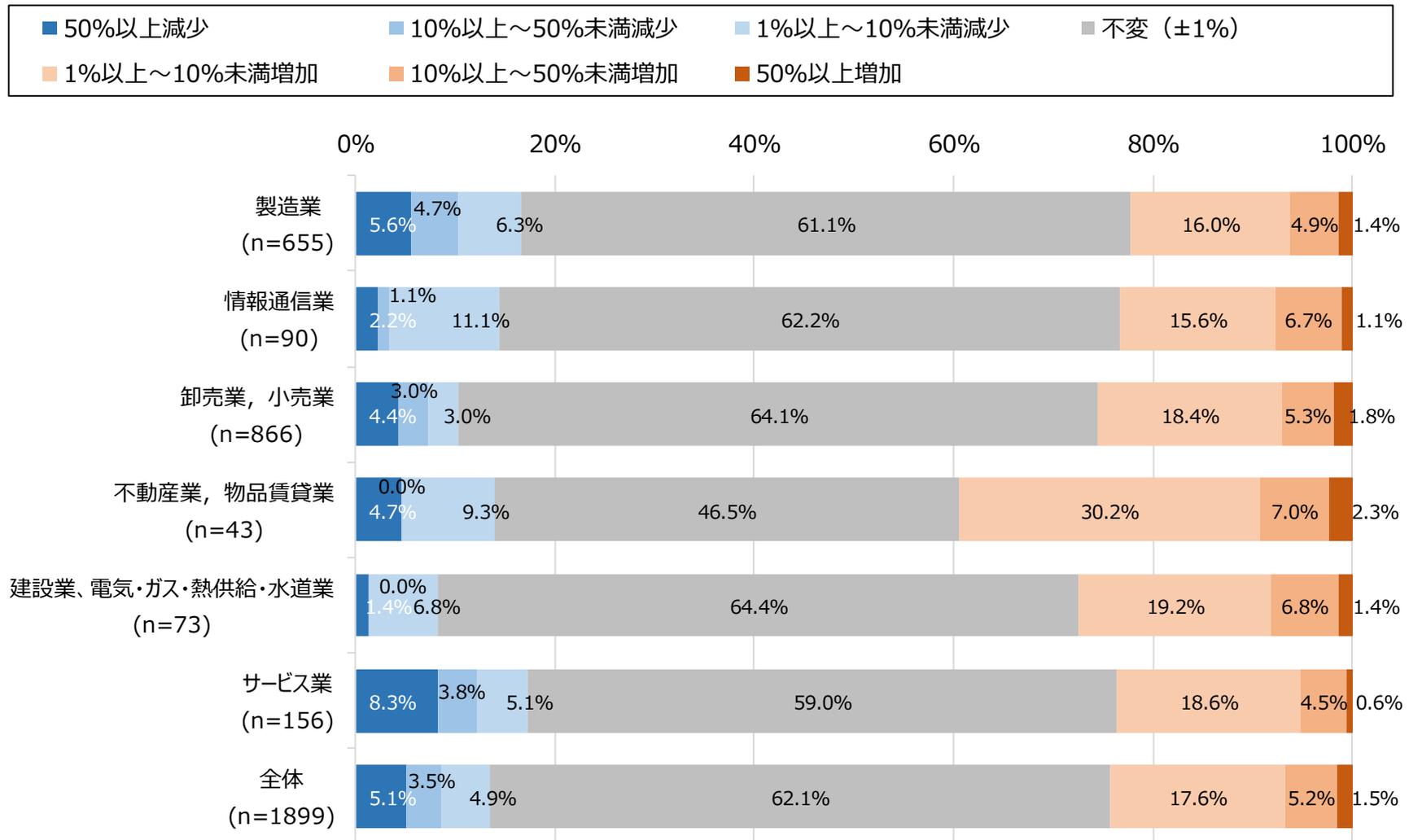
Q17：新型コロナウイルス感染症の影響により、貴社の経営等にどのような影響（変化）がありましたか？2019年度と比べて2020年度がどのように変化したかをお答えください。

※売上高



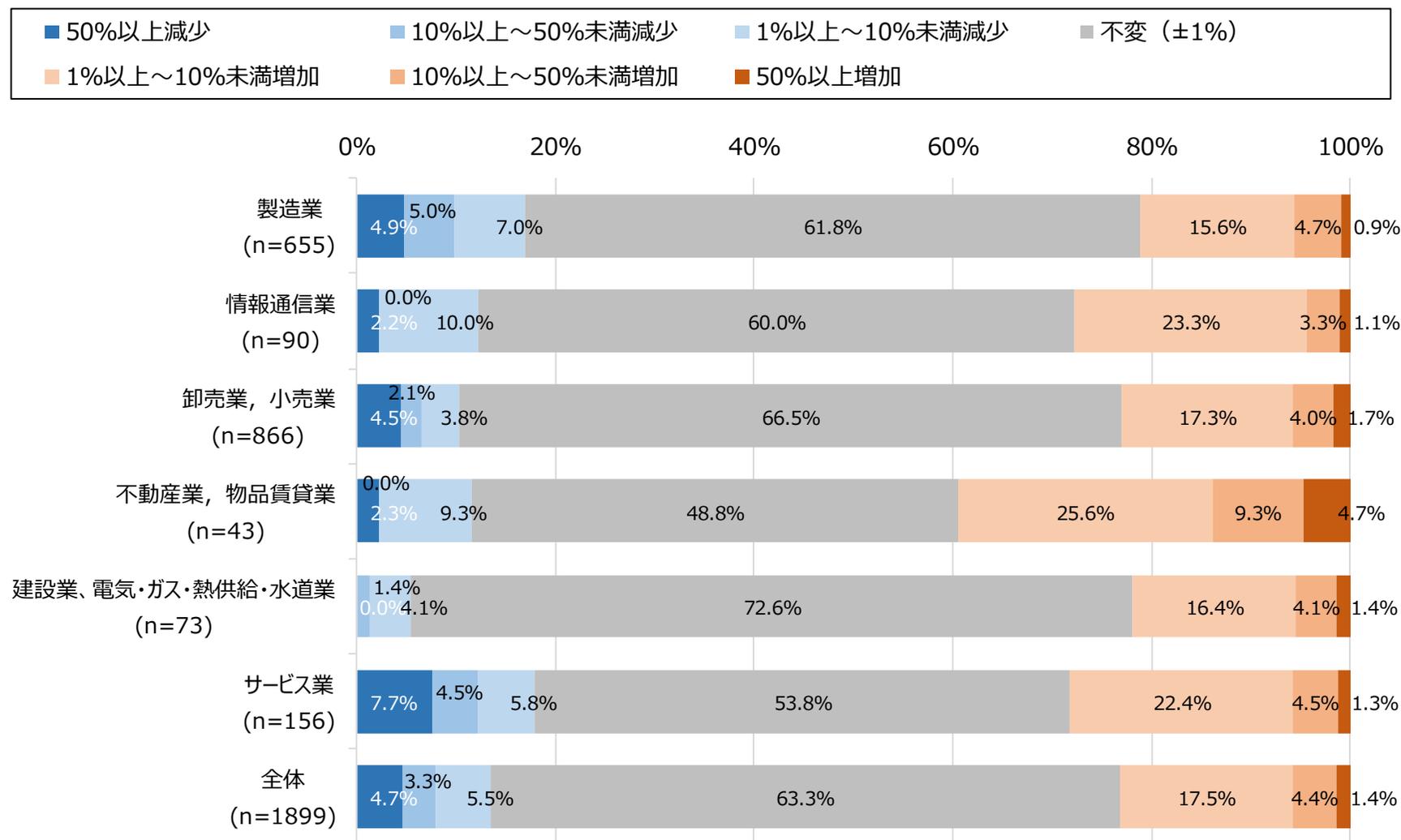
Q17：新型コロナウイルス感染症の影響により、貴社の経営等にどのような影響（変化）がありましたか？2019年度と比べて2020年度がどのように変化したかをお答えください。

※ICTハードウェア投資



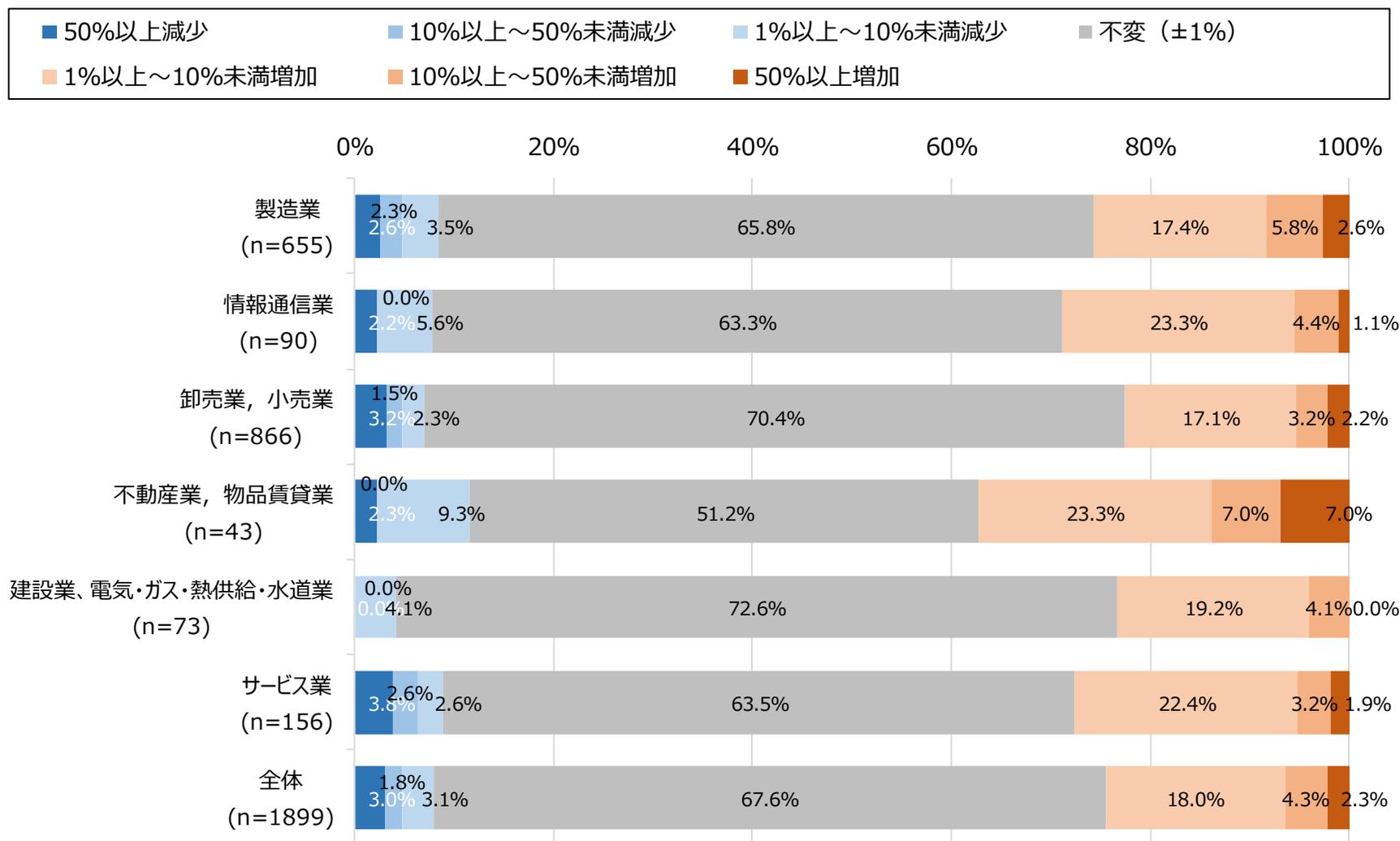
Q17：新型コロナウイルス感染症の影響により、貴社の経営等にどのような影響（変化）がありましたか？2019年度と比べて2020年度がどのように変化したかをお答えください。

※ICTソフトウェア投資



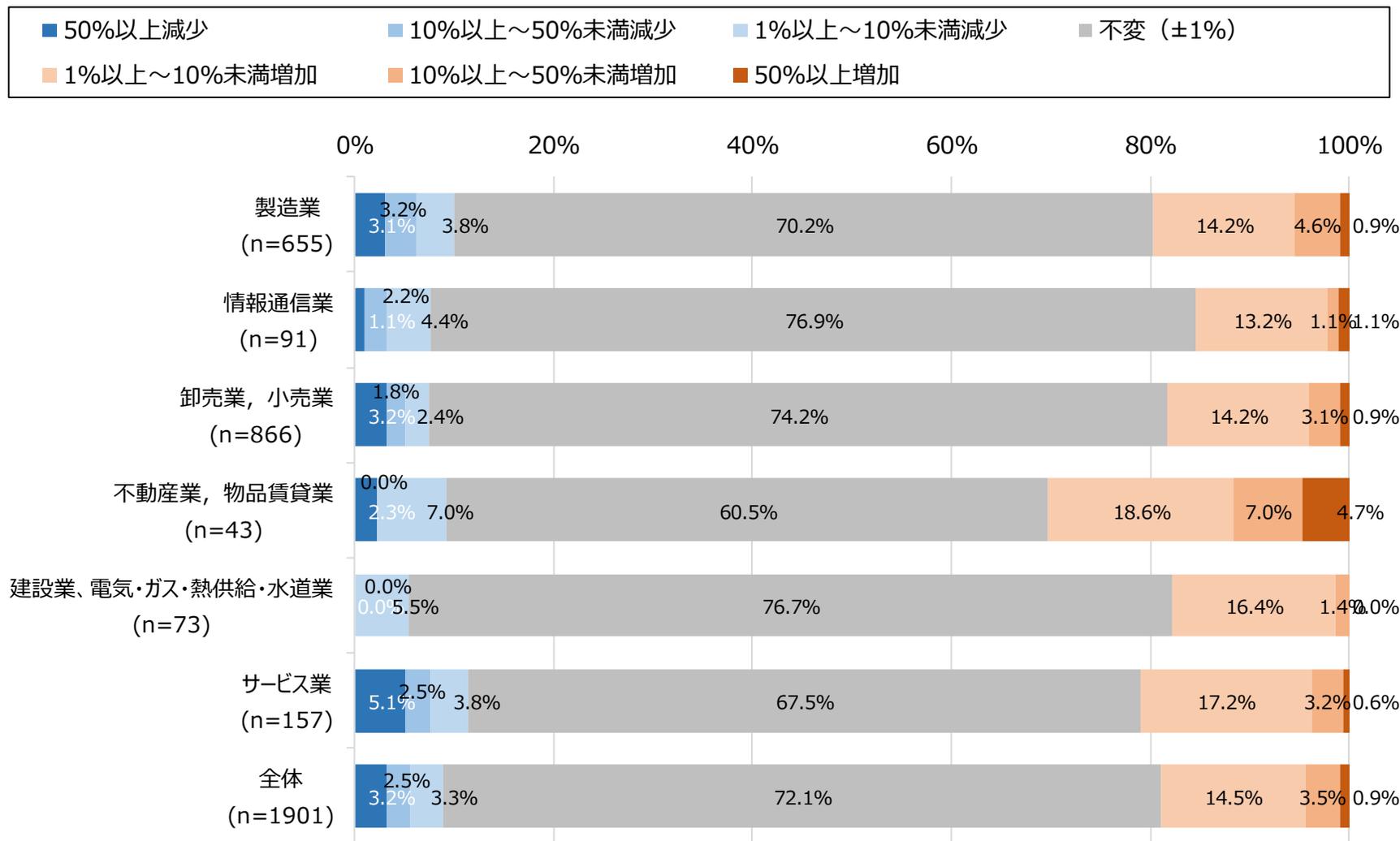
Q17：新型コロナウイルス感染症の影響により、貴社の経営等にどのような影響（変化）がありましたか？2019年度と比べて2020年度がどのように変化したかをお答えください。

※クラウドサービス支出

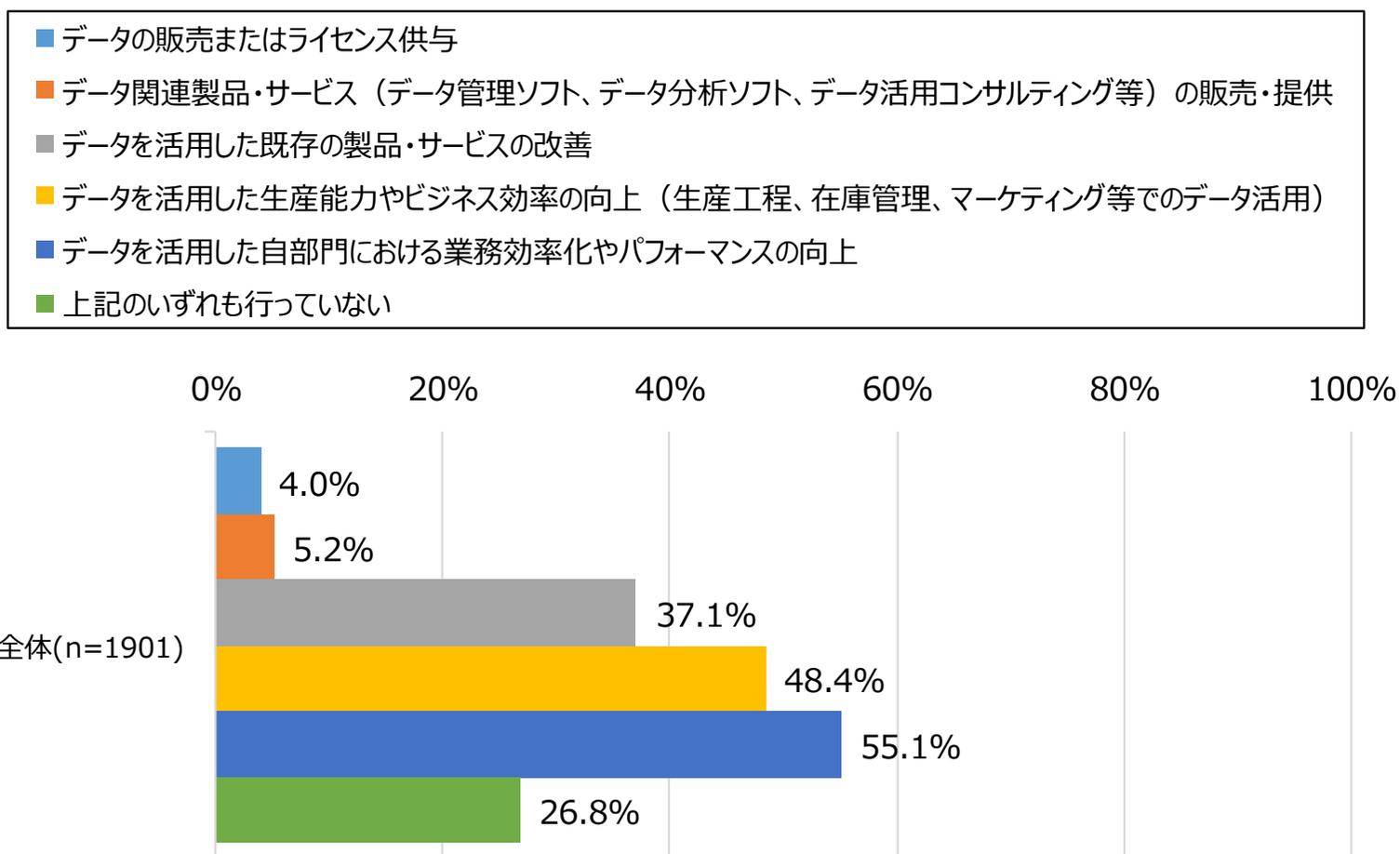


Q17：新型コロナウイルス感染症の影響により、貴社の経営等にどのような影響（変化）がありましたか？2019年度と比べて2020年度がどのように変化したかをお答えください。

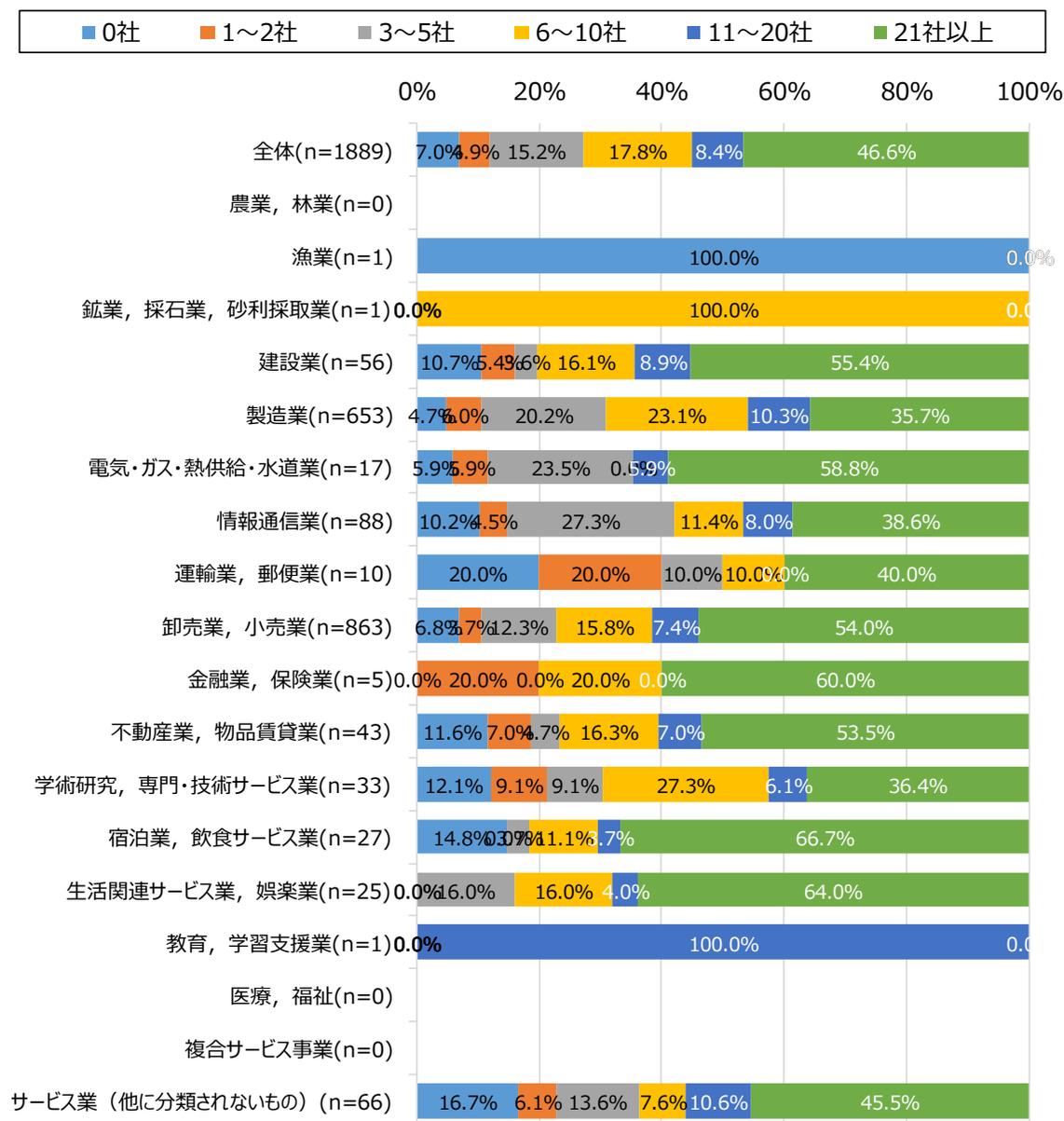
※データ活用に関連する投資・支出



Q18：貴社では、データを収益につなげる取組として、どのようなものを手がけていますか？



Q19：貴社の代表的な製品・サービスを投入している市場において、2019年度末時点で、貴社と競合する企業は国内外合わせて何社程度ありましたか？



Q20：貴社では、どのような目的でデータ活用に取り組んでいますか？自由にお答えください。

※全企業回答（任意）

<抜粋>

- お客さまに支持していただける品揃え・在庫を実現する。後方作業を削減する。
- お客様満足、ロス対策、売上対策
- バックオフィス業務の効率化に今まで重点を置いていたが、マーケティングへの活用を拡大させるのが今後の課題
- マーケティング方針の妥当性検証 業務の効率化
- 業務の効率化 在庫の削減
- 業務効率化及び経営判断の為
- 現在は、事務（処理）の効率アップの目的でのみ、データ活用を行っています。「攻め」の施策に絡んだデータ活用は未実現の状況です
- 個々のナレッジ共有によるコミュニケーション強化と付加価値向上
- 顧客の販売動向分析およびサービスの提供 顧客の地域・購買層の把握
- 購買動向の把握と新規ブランド・商品の投入判断
- 市場（顧客）の動向把握のため。将来力を入れる市場を予測するため 当社の弱点を把握し改善点を顕在化するため。
- 市場分析(製品、エリア、顧客等の戦略)
- 社内各セクションの実績把握及び検証
- 新たなサービスの開発、販促費支出の効率化、意思決定の迅速化
- 販売結果を見て、次の施策に反映させる
- 目的は、①収益向上、②業務の効率化、③リスク低減
- 与信管理、クレーム処理、品質保証 製品開発、販売促進
- 利益率の向上を最大ポイントとして利益率の低い顧客を抽出して対策を打っています。場合によっては取引を停止します。
- 目先は業務の効率向上や生産性向上を目的としており、将来はデジタル技術を活用した新たな付加価値を生み出せることを期待して取り組んでいる。

Q21：貴社では、データ活用に関連してどのような課題をお持ちですか？

人材・スキル・システム・データ不足といった組織内の課題やデータの外部共有に関する障壁など自由にお答えください。

※全企業回答（任意）

<抜粋>

- 時系列でデータを探ることが困難（経済センサスなど使い勝手がよくない）。必要なデータを探ることができる人材が少ない。社内にどのようなデータがあるか共有できていない。など
- データ入手の費用対効果が不明で入手に踏み切れない。社内でのデータの分散。データ分析に係る社内コミュニケーション
- データ分析人材およびスキル不足。必要なデータ不足。データ連携するキー不足。コード体系
- ICTに対するノウハウとアナリストとしてのノウハウの両方を兼ね備える人材
- クラウド系インフラへの移行
- システムが用途ごとにバラバラなため、データ連携や分析は原則行っておらず、必要な場合は手集計で行っているため効率化を図りたい
- スキルもリソースも指針も無い。
- データーを統合的に活用できる人材不足
- データのインプットの手間の簡略化
- データをどういう形で企業活動に生かしていくのか、目的や目標が明確になっていないように感じている。
- 外部とのデータ連携の具体化
- 自社独自開発システムと外部システムとの整合性
- 失敗を許容する企業文化
- 人材・スキル・システム・データ不足
- 当社で使用するシステムに統一性(互換性)がとれていないものがある。将来的には全システムがデータ共有できるものに変えていきたい
- 入荷データがデジタル化していないため、毎日手入力が全体の60%を占め非効率である。
- 人材不足 経営層にITに強い人材がいない
- 社内業務プロセスが属人化、サイロ化して情報が全社で一元化出来ない（Excel多用など）。これを改革、断行する作戦、具体化が弱い（スキル面、体制面などで）。
- 社内IT管理者の不足、営業・事務部門のITスキルの向上、データ分析専任者の育成 に関わるコスト増
- データ分析業務にあたる者が、他の業務も兼務しているため、思うように時間が取れない
- データの意匠権侵害への脅威。

2. 関係性分析
3. 実証分析結果
4. ストーリーに基づいた分析

【構成員限り】

(参考) 昨期のアンケート調査項目と今年度の対応①

今期は、昨期のアンケート調査項目を元に、データの活用状況をより適切に把握する観点や調査の継続性を念頭に調査項目を設定した。昨期との比較は下記のとおり。

アンケート調査項目	昨期の調査票設問	今年度の対応
(Q1)データの分析手法 [種類別×領域別]	貴社では、どのようなデータをどのように処理したものを各領域で活用していますか？当てはまるものすべてをお答えください。	3重クロスの設定で回答者負担が大きいため簡易化
(Q2)分析に活用するデータの量(期間) [種類別]	貴社では、データ分析を行う際、おおよそどれくらい前のデータまでを利用しますか？	昨期の調査結果を踏まえて選択肢を変更
(Q3)分析の頻度 [領域別]	貴社では、データ分析をどのくらいの頻度(間隔)で行いますか？	昨期の調査結果を踏まえて選択肢を変更
(Q4)データの品質に関連した取り組み	貴社ではデータを活用する際に、データの品質に関連したどのような取り組みを行っていますか？	削除
(Q5)分析の体制、データ分析の専門部署が立ち上がった時期	貴社がデータ分析を行う際、どのような体制(部署・担当者)で分析(外部企業等が提供しているデータ解析ツール・サービスの利用も含みます)を行いますか？データ分析を行う専門部署がある場合は、どのくらい前に当該部署が立ち上がったのかについてもお答えください。	設問を分割し、分析の体制についてのみ確認
(Q6)分析の人員数と学歴 ※現在、3年前	貴社で、データ分析を主な業務とする従業員(正社員だけでなく、データ分析に寄与するパートタイマー・アルバイト・契約社員・派遣社員等を含みます)はおおよそ何人くらいいますか？ また、その内、大学院卒(修士または博士の学位を有する)の学歴をもつ従業員の割合はどの程度ですか？	設問を分割し、人員数についてのみ確認
(Q7)分析の手法(AIの活用状況)	貴社では、データ活用において、どのようなAI技術を用いていますか？	削除
(Q8)データ活用の各プロセスにおける効果への影響	データを活用することによる効果(企業活動に対する貢献)についてお伺いします。貴社においては、データ活用に関わる各プロセスが効果にどの程度貢献していると思いますか。	削除
(Q9)データの多様性(データの入手元、提供状況) [種類別]	貴社で蓄積(保有)されているデータについて、データの入手元はどちらですか？また、保有しているデータを必要に応じて加工した上で、他社等に提供または共有していますか？	設問形式を変更し、入手元の種類、国内・国外、有償・無償別に確認
(Q10)外部から購入したデータの割合	貴社で蓄積(保有)しているデータの内、外部から購入したデータ(無料で公開されているデータやアライアンス等による無償の共同利用のデータは含みません)はおおむねどの程度を占めていますか？データの容量ではなく、件数ベースでお答えください。	外部から購入したデータについて支出額を確認する形に変更

(参考) 昨期のアンケート調査項目と今年度の対応②

アンケート調査項目	昨期の調査票設問	今年度の対応
(Q11)データの蓄積量（容量割合、件数）[種類別]、データの総容量 ※現在	<p>貴社で蓄積（保有）されているデータについて、2018年度末時点におけるおおよその構成割合（容量ベース）を合計が100%になる形でお答えください。</p> <p>また、蓄積（保有）データの対象数（人・社・者・台数）を、有効数字2桁の概数（例：12,000、450,000）でお答えください。対象数が不明な項目については、お答えいただかなくても構いません。併せて、蓄積（保有）されている具体的なデータの種類について可能な範囲でお答え下さい。</p> <p>上記でお答えいただいたデータについて、2018年度末時点における容量の総計を有効数字2桁の概数（例：12,000、450,000）でお答えください。正確な値を把握されていない場合は、想定される範囲の中間値（例：1,000～2,000TB程度と思われる場合は1,500TB）をお答えください。</p>	<p>設問を分割し、活用しているデータ容量と直近1年間の増減について確認。</p> <p>また、活用しているデータの容量割合についても確認</p>
(Q12)データの蓄積量（自社以外から購入・入手したデータの件数割合）[種類別] ※現在	<p>貴社で蓄積（保有）されているデータについて、2018年度末時点で自社以外から購入・入手したデータ（無料で公開されているデータやアライアンス等による無償の共同利用のデータも含みます）は蓄積されているデータのどれくらいの割合を占めていますか？データの容量ではなく、件数ベースでお答えください。</p>	<p>データの件数割合ではなく、容量割合に変更</p>
(Q13)データの蓄積量（件数）[種類別] ※3年前からの変化	<p>貴社で蓄積（保有）されているデータについて、2015年度末時点と比べて2018年度末時点のデータはどの程度変化（増減）しましたか？データの容量ではなく、件数ベースでお答えください。</p>	<p>削除</p>
(Q14)データの蓄積量（非構造化データの件数割合）[種類別]	<p>貴社で蓄積（保有）されているデータについて、非構造化データ（業務日誌やSNSの書き込み、音声/画像/映像など規則（構造）が明確に定義されていないデータ）はどれくらいの割合を占めていますか？データの容量ではなく、件数ベースでお答えください。</p>	<p>削除</p>
(Q15)分析に活用するデータの件数割合 ※現在、3年前	<p>貴社で蓄積（保有）している全データの内、1年間（2018年度）に実際にデータ活用に用いたデータの割合はおおむねどの程度ですか？データの容量ではなく、件数ベースでお答えください。また、3年前（2015年度）時点の状況と合わせてお答えください。</p>	<p>削除</p>
(Q16)新製品・新サービスの投入数	<p>貴社では2018年度に新製品・新サービスをどれくらい市場に投入しましたか？</p>	<p>削除</p>
(Q17)競合数 ※現在、3年前からの変化	<p>貴社の代表的な製品・サービスを投入している市場において、2018年度末時点で、貴社と競合する企業は国内外合わせて何社程度ありましたか？また2015年度末時点から2018年度末時点の変化についても合わせてお答えください。</p>	<p>そのまま継続採用</p>

企業が付加価値を生み出す生産要素として「資本」、「労働」及び「データ」を位置づけ、一次同次を仮定しない生産関数を推定。

$$V = A_0 K^\alpha L^\beta Data^\gamma, \quad \log(V) = \log A_0 + \alpha \log(K) + \beta \log(L) + \gamma \log(Data) + \text{業種ダミー}$$

※製造業、非製造業に分類

実証分析の結果から、活用データ容量・件数は、他の生産要素（資本、労働）と同様に、付加価値に対してプラスの関係性を持っていること、また、関係性の程度は、活用データ容量・件数が1%増えると付加価値が0.05%増える程度であることがわかった。

データ変数	n数	修正R ²	K (資本)	L (労働)	Data (データ)
活用データ容量 (=データ総容量×分析に活用するデータの割合)	258	0.8343	0.44 ◎	0.50 ◎	0.05 ◎
活用データ件数 (=データ総件数×分析に活用するデータの割合)	135	0.8157	0.34 ◎	0.55 ◎	0.07 ◎
活用データ容量 (2015年度)	258	0.8332	0.44 ◎	0.51 ◎	0.05 ◎
活用データ件数 (2015年度)	135	0.8136	0.34 ◎	0.56 ◎	0.06 ○
外部入手データ容量 (=データ総容量×外部から入手したデータの割合)	267	0.8401	0.47 ◎	0.47 ◎	0.05 ◎
外部入手データ件数 (=データ総件数×外部から入手したデータの割合)	140	0.8193	0.35 ◎	0.51 ◎	0.07 ◎
内部保有データ容量 (=データ総容量 - 外部入手データ容量)	267	0.8379	0.47 ◎	0.48 ◎	0.05 ○
内部保有データ件数 (=データ総件数 - 外部入手データ件数)	140	0.8134	0.35 ◎	0.55 ◎	0.06 ○
データ総容量×データ活用度 (活用領域・種類・処理方法の多さ)	261	0.8334	0.44 ◎	0.49 ◎	0.05 ◎
データ総件数×データ活用度	137	0.8150	0.32 ◎	0.58 ◎	0.05 ○
データ総容量×データの多様性 (データの入手・提供の多様性)	173	0.8537	0.53 ◎	0.42 ◎	0.04 ○
データ総件数×データの多様性	85	0.8576	0.41 ◎	0.49 ◎	0.06 ◎

(注) ◎:有意水準1%、○:有意水準5%、△:有意水準10%

※2015年度と記載のない変数は2018年度の値、外部から入手したデータの割合は、データの外部入手状況を踏まえて補正を行った。

- データの活用プロセスは収集、加工、分析、解釈（活用）に整理できる。そこで、データ活用の成功要因をそれぞれのプロセスにおける社内リソース、社外リソース、活用技術、関連する取組、その他で整理した。

	収集	加工	分析	解釈（活用）
社内リソース	<p>(A社) ホテル宿泊する際のアンケートは全て自社で行っており、現在は全てウェブで回答してもらっている。回収率はインセンティブなしで30%を目指しており、30%に行っている施設もあり、かなり回収できている。宿泊客が登録したら、分析するシステムに反映され、スタッフがリアルタイムに見ることができる。</p> <p>(B社) 予測モデルの精度は説明変数に依存する。欲しいデータをいかに収集・加工するかが大事。</p>		<p>(D社) データ分析は大学と連携している。開発と研究は内製化をしている。大学の先生に意見をもらっている。社内にAI技術者がいる。AI自体が新しいものなので、元々AIの知識がある方ではあるが、学びながら取り組んでいる。</p>	<p>(B社) データサイエンティストだけいてもうまくいかない。財務、人事など、データの内容がわかる人が必要だ。</p>
社外リソース	<p>(C社) 病院のデータ活用の目的は、患者の情報を集めて、患者ごとのサービスの充実化として活用しようとしている。</p>			
活用技術	<p>(C社) 病院はまだクラウドに懸念を抱いている。実際に稼働しているシステムにクラウドはほとんどない。</p>	<p>(D社) AI分析で一番時間かかるのは、教師データの作成だ。そこで、クリックただけで、簡単に作れるシステムを開発した。そのシステムを利用することで、未経験者でもデータを作成できる。</p>		<p>(F社) 将来的には需要予測の精緻化を考えている。在庫の削減、需要見込みを見誤ることによるロスの低減、計画的に配置できることによる車両台数、運搬コストの低減が期待できる。</p>
関連する取組	<p>(A社) CSデータや獲得データ、宿泊客の施設での滞在データなどが有機的につながっていない。</p>	<p>(D社) 今後、画像データが3Dに変わる時に、データがビッグデータ化していくので、クラウドでの保存等の方法を考えないといけない（容量が増え、かつAIなどに活用する必要がある）。</p>		<p>(G社) 1つはお客様へのレポート提供によって、旅の価値を高める。さらに外部のレコメンデーションサービスを紹介する。</p>
その他	<p>(B社) 課題を定義することが大事。課題を定義して、必要なデータを集め、アジャイル的に試行錯誤する。</p>	<p>(D社) 10年前のデータはCADの図面化に活用できるかもしれないが、AIのデータは取り方など細かく決まっているので、昔すぎるデータはAIに使えないかもしれない。</p>	<p>(E社) データの価値が継続する期間について、10年前のデータも現在のデータも高齢者の生体情報についてはあまり価値は変わらないだろう。</p>	<p>(A社) マイクロツーリズムのキャンペーンにターゲットを絞ってアプローチする際に活用した。持っているデータももう少し効率的にマーケティングに活用できるだろうと思う。そこはまだ課題がある。</p>