

データ分析とプレゼンテーション技法

科目コード

AB1035



単位数	履修方法	配当年次	担当教員
2	R or SR(講義)	1年以上	岩田 一樹

科目の概要

■科目の内容

本科目では、「データの活用」に主眼をおき、実生活や職場で実践できる「データ分析の基礎」を学びます。そのため、統計学、および、数学的な内容については必要最低限の範囲に留め、その代わりに、「データ分析の設計方法」、「データのクレンジング」、「分析結果の解釈と提示時の留意点」等のより実践的かつ実務的なデータ分析に関係する事柄について学びます。

なお、この科目ではデータ分析および分析結果の提示等に Excel や PowerPoint を使用しますので、Microsoft 社の Office がインストールされているパーソナルコンピュータが必要になります。

【スクーリングで学ぶ内容】

スクーリングでは、データ分析に関する基本的なプロセスを学び、データ分析を演習形式で実際に行います。

【レポート学習で学ぶ内容】

レポート学習では、データ分析における基本的な語句の意味についてとそれらの計算、および、問題とデータが与えられた際のデータ分析のプロセスについて学びます。

■到達目標

本科目では、以下の6点を到達目標とします。

- 1) 「データ分析」の目的を適切に説明できる
- 2) データ分析を実施する上で重要な「問題領域」「評価軸」「要因」について説明でき、それらを立案できる
- 3) データのグラフ化(可視化)を行い、データの概要(特徴等)の把握ができる
- 4) 「代表値」(平均値、中央値、最頻度)の意味と使う場面を説明できる
- 5) データ分析の結果を他者に正しく伝えるためのグラフ化ができる
- 6) PowerPoint を用いてデータ分析および分析結果をプレゼンテーションできる

■教科書

河村真一・日置孝一・野寺 綾・西脇清行・山本華世著『本物のデータ分析力が身に付く本』日経BP社、2016年

(スクーリング時の教科書) 上記教科書は必ず持参してください。

■「卒業までに身につけてほしい力」との関連

本講義を受講することで、

- ・データの収集・分析のみでなく、データ分析の設計までを包含した「ICT 活用力」
- ・問題設定、評価軸、評価要因の設定、データの分析、結果の解釈・提示を通じた「問題解決能力」
- ・データ分析結果の解釈を通じた「自己コントロール力」
- ・全体を通して、データ分析による「社会貢献力」

を身につけていただけます。

■科目評価基準

レポート評価40%+スクーリング評価 or 科目修了試験60%

■参考図書

東京大学教養学部統計教室編『統計学入門』東京大学出版会、1991年

R. Ennos 著『パソコンで簡単！すぐできる生物統計—統計学の考え方から統計ソフト SPSS の使い方まで』羊土社、2007年

柏木吉基著『データ・統計分析が出来る本』日本実業出版社、2013年

玄場公規・湊 宣明・豊田裕貴著『Excel で学ぶ ビジネスデータ分析の基礎』オデッセイコミュニケーションズ、2016年

門脇香菜子著『できる Excel ピボットテーブル データ集計・分析に役立つ本』インプレス、2016年

■事前に学習してほしい科目（任意）

「情報処理の基礎」、または文章作成、表計算の基本的な技術の修得をしていること

スクーリング

■スクーリング受講にあたっての留意事項

原則としてスクーリング申込締切日までに、1単位め課題を提出してください。受講者多数で受講定員を超える場合、提出のない方は受講できません。

■講義内容

回数	テーマ	内容
1	データ分析の設計	データ分析の目的、基本的な内容、設計方法について学びます
2	データのクレンジング	実際のデータを用いて「外れ値」、「欠損値」の確認方法、および、データのクレンジングについて学びます
3	データ分析方法の選択	「代表値」と「クロス集計」について、その意味と方法を学びます
4	標準偏差	標準偏差の意味、標準偏差の使い方、標準偏差の計算方法について学びます

回数	テーマ	内容
5	グループ間の差の検証	グループ間の比較を行う際の注意点とその判断方法について学びます
6	分析結果の捉え方	分析結果の解釈、および、表現について学びます
7	プレゼンテーションの作成	分析結果を効果的にプレゼンテーションする方法を学びます
8	質疑応答	
9	スクーリング試験	

■講義の進め方

スクーリングではパワーポイントによる説明、および、グループワークを行います。

■スクーリング 評価基準

スクーリング試験では、データ分析の設計、および、各用語の意味とその適切な使用方法の理解について問います。(教科書、配布資料、自筆ノート 持ち込み可)

履修者数によっては、スクーリングで実施するデータ分析の報告をもって試験とする場合があります。

■スクーリング事前学習 (学習時間の目安：5～10時間)

講義内容で関心あるテーマについて、自分なりに学びたいことを考えてきてください。

なお、原則としてスクーリング申込締切日までに、1単位め課題を提出してください。

■スクーリング事後学習 (学習時間の目安：20～25時間)

興味のあるテーマに関するデータについて、自分なりに分析に取り組んでください。その際は可能な限り、考えたこと、行ったことについてノートを取るようになしてください。

レポート学習

■在宅学習15のポイント

回数	テーマ	学習内容・キーワード	学びのポイント
1	データ分析の設計①問題領域の決定	学習内容 問題領域の設定について学びます	新しい“問い”に対してデータ分析を行う意味と、その問いに対しての問題領域とは何かを理解してください。
2	データ分析の設計②評価軸と要因	学習内容 1で学んだ問題領域についての評価軸とその評価のための要因について学びます	設定された問題領域に対して、その評価に関することを理解してください。
3	データ分析の設計③分析の概念図の作成	学習内容 1、2で学んだことを併せてデータ分析の概念図の作成について学びます	1、2で学んだことを系統だててまとめ上げるとともに、可視化することの有効性について学んでください。

回数	テーマ	学習内容・キーワード	学びのポイント
4	データクレンジング① 出自と概要の確認	学習内容 データ分析を行うにあたり、データの出自を確認することの重要性を学びます キーワード：可視化	データ分析をする際に、対象のデータの出自を明確にすること、および、分析前にデータの概要を把握しておくことの重要性を理解してください。また、Excelを用いたデータの可視化について修得してください。
5	データクレンジング② 外れ値と欠損の取扱	学習内容 データにおける外れ値と欠損値の取扱について学びます キーワード：外れ値、欠損値、クレンジング	データの外れ値と欠損値とは何かと、それらの扱い方について理解してください。これらの取扱いはクレンジングと呼ばれ、分析結果に大きな影響を与える場合があるので、大変重要なものになります。
6	分析方法の選択①代 表値	学習内容 代表値である平均値、中央値、最頻度について学びます。 キーワード：平均値、中央値、最頻度	それぞれの代表値の意味すること、および、4で学んだデータの概要に対してどの代表値を使用するのが適切なかを理解してください。
7	分析方法の選択②ク ロス集計	学習内容 データ分析にあたりクロス集計を行うことの有効性を学びます キーワード：クロス集計、ピボットテーブル	データ分析において重要な手法であるクロス集計について、その有用性、および、Excelによる実施方法について修得してください。
8	標準偏差①標準偏差 とは何か？	学習内容 標準偏差についてとその算出方法について学びます キーワード：標準偏差、母集団、標本集団	データにおける標準偏差とは何かについて理解し、Excelを用いたその算出方法について理解してください。
9	標準偏差②標準偏差 の利用	学習内容 8で学んだ標準偏差についてその利用方法について学びます	8で修得した標準偏差がどの様にデータ分析に活かされるのかを、可視化されたデータと照らし合わせながら理解してください。
10	グループ間の差の検 証①検証の考え方	学習内容 7で学んだクロス集計に関連して、そのグループが2種類の場合に、そのグループ間の違いをどう検証するかを学びます キーワード：確率、検定	2グループ間の差について客観的な根拠を与えるには確率の概念を用います。ここでは、その考え方について理解してください。

回数	テーマ	学習内容・キーワード	学びのポイント
11	グループ間の差の検証②検証結果と判断	学習内容 10で学んだ検証の考え方を元に、検証の行い方とその結果の解釈の仕方を学びます キーワード：検定	10に続いて、2グループ間の差を検証する具体的な Excel を用いた手法を修得してください。 (なお、3グループ以上間の比較には「分散分析」と呼ばれる手法を用いますが、本科目の範囲を超えています。興味のある方は、参考図書『統計学入門』をご参照ください。)
12	結果の解釈	学習内容 データ分析結果について、その結果の解釈を行う際の注意点について学びます キーワード：確証バイアス	データ分析の結果を解釈・判断するのは人間です。そこには、思い込みなどの結果の解釈・判断を狂わせる可能性のあるものが幾つかあります。判断を狂わせる代表的なものを把握し、適切な判断を下せる確率を上げてください。
13	結果の他者への伝え方	学習内容 データ分析結果について、その結果を提示する際の注意点について学びます キーワード：可視化	データ分析の結果は図表で提示することが多いですが、その際の注意点について学んでください。
14	ケース実習①データ分析の設計	学習内容 1～3までに学んだ内容を基に実際にデータ分析の設計を行います	プロセスは実際に実施して習得することができます。教科書4章に沿って、是非、ご自身で手を動かしてデータ分析を実施してください。
15	ケース実習②データ分析の実施と結果の提示	学習内容 4～13までに学んだ内容を基に、14で設計したデータ分析についてデータ分析を行います	14に引き続き、手を動かしてデータ分析を体験してください。

■レポート課題

1 単位め	データ分析に関する基本的な語句、その意味、その算出について「客観式レポート」の問題を解答してください。																									
2 単位め	<p>あなたは某フードチェーンに勤務していて、新店舗の出店企画を任されました。出店候補地としては「駅前」、「住宅地」、「郊外」の3カ所があり、下記のデータが与えられています。この与えられたデータのみから、あなたなら上記3カ所の「何処」に出店するか決断し、その理由と併せて解答してください。</p> <p>なお、出店を行うことは決定しており「出店を行わない。」という選択肢はありません。また、説明には図表を用いてください。</p> <p>※下記データは、通信教育部ホームページ『レポート課題集C（心理専門編）2019』からエクセル表をダウンロードすることが可能です。ダウンロードできない方は、メールでご相談ください (ur@tfu-mail.tfu.ac.jp)。</p> <table border="1"> <thead> <tr> <th>店舗番号</th> <th>立地（1: 駅前、2: 住宅地、3: 郊外）</th> <th>売り上げ（円 / 日）</th> <th>来客数（人 / 日）</th> <th>客単価（円 / 人）</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>3</td> <td>181,056</td> <td>873</td> <td>207</td> </tr> <tr> <td>2</td> <td>1</td> <td>282,486</td> <td>694</td> <td>407</td> </tr> <tr> <td>3</td> <td>3</td> <td>195,752</td> <td></td> <td></td> </tr> <tr> <td>4</td> <td>1</td> <td>260,198</td> <td>837</td> <td>311</td> </tr> </tbody> </table>	店舗番号	立地（1: 駅前、2: 住宅地、3: 郊外）	売り上げ（円 / 日）	来客数（人 / 日）	客単価（円 / 人）	1	3	181,056	873	207	2	1	282,486	694	407	3	3	195,752			4	1	260,198	837	311
店舗番号	立地（1: 駅前、2: 住宅地、3: 郊外）	売り上げ（円 / 日）	来客数（人 / 日）	客単価（円 / 人）																						
1	3	181,056	873	207																						
2	1	282,486	694	407																						
3	3	195,752																								
4	1	260,198	837	311																						

2 単位め

店舗番号	立地 (1: 駅前、2: 住宅地、3: 郊外)	売り上げ (円 / 日)	来客数 (人 / 日)	客単価 (円 / 人)
5	1	215,480	466	462
6	2	261,165	1,075	243
7	2	209,642	492	426
8	2	268,757	784	343
9	3	276,490	498	555
10	1		931	
11	1	305,940	107	2,859
12	2	277,052	919	301
13	3	202,473	285	710
14	2	195,214	899	217
15	3	230,603	703	328
16	2	276,625	182	1,520
17	2	196,495	512	384
18	3	246,167	421	585
19	2	195,041	612	319
20	3		842	
21	1	294,140	908	324
22	3	191,419	326	587
23	1	295,571	827	357
24	2	252,784	548	461
25	2	205,250	553	371
26	3	208,434	704	296
27	2	229,499	737	311
28	3	187,434	656	286
29	3	216,179	653	331
30	1	293,021	695	422
31	1	273,057	154	1,773
32	2	249,442	392	636
33	3	261,922	927	283
34	3	183,260	305	601
35	2	247,642	902	275
36	1	220,597	231	955
37	1	295,889		
38	1	245,454	305	805
39	1	253,360	978	259
40	2	252,911	934	271
41	3	277,342	468	593
42	3	214,971	688	312
43	2	222,813		
44	3	190,219	674	282
45	3	271,821	870	312
46	3	271,264	1,042	260
47	2	281,536	767	367
48	2	256,585	989	259
49	1	223,302		
50	1	261,001	693	377
51	2	254,378	190	1,339

※提出されたレポートは添削指導を行い返却します。

全体

受講生されるみなさんは「データ」と言われると何を思い浮かべるでしょうか？「国勢調査」や「〇〇白書」等の統計調査の結果、「野球の打率や防御率」等の「数字」を思い浮かべるかもしれませんが、「データ」はこれだけではありません。「データ」の身近な具体例を示すと、みなさんの携帯電話や電子メールによるやり取り、みなさんも書かれているかもしれないブログの内容、Twitterの眩き、Facebook等の「SNS」(Social Networking Service)を介してのネットワークコミュニティ等が挙げられます。これらは、先に挙げた「国勢調査」等とは別物のよう感じられるかもしれませんが、実は、これらは全て何らかの手法で「数字」に置き換えることが可能なのです。例えば、Twitter上のつぶやきは、使用されている各単語に番号を付けて数値化します。つまり、「データ」とは「数字」で表現される何らかの意味(人に役立つ知識)を包含したもののなのです。

そして、上記で列挙した例をご覧になって、それらの多くがインターネットと深く関連していると思われる方がいらっしゃるかもしれませんが、それは間違いではありません。そのインターネットの発展によるデータ取得の容易化を背景に、近年、とくに2010年頃から、世界的に「ビッグデータ」と呼ばれる大量データを活用してビジネスや世の中に役立てようという風潮が高まっています。受講生のみなさんも「ビッグデータ」や「IoT」(Internet of Things =もののインターネット)といった語句をご覧になったことがあると思います。

では、本科目の目的でもある「データの活用」とは何なのでしょう？それは、「人の意思決定(新たな“問い”に答えるの)に際して、適切な答えを導き出すのに数字(データ)を活かす」ことに他なりません。そして、その決定に対して数字(データ)から意思決定に役立つ「新しい知識を引き出す」ことこそが「データ分析」なのです。(データから新たな知識を見出すことは「データ・マイニング」と呼ばれることもあります。)

ここで、一点注意すべきことがあります。それは、「データ分析」とは「数字(データ)から意思決定に役立つ新しい知識を引き出す」こと、と述べましたが、引き出された「知識」が正しいかどうかの評価は誰にも出来ないという点です。このことは、純粋無垢な原理から公理を導き完璧な「結果」が存在する数学とは異なり、「データ分析」では「現実世界」を数字に置き換えたものから知識を導くことに起因します。すなわち、現実の世界を数字に置換する際に誤り(誤差)や欠損等を包含してしまう完璧でない「データ」から知識を導く「データ分析」においては、絶対に正しい完璧な結果というものは存在しないということです。そして、その絶対に正しいとは保証されない知識(結果)を用いて行う意思決定も、必然的に、その決定が絶対に正しいということはありません。このような理由から「データ分析」および「意思決定」は導かれた「結果」での評価ができないために、これらの「正しさ」は「結果」そのものではなく、その「結果」を導いた「プロセス」によって評価されることとなります。その具体的なプロセスとは「データ分析の設計」、「データの事前チェック」、「分析方法の選択」、「分析結果の評価・解釈」、「分析結果の表現」です。したがって、正しい「データ分析」を行うためにはこのプロセスを習得すれば良いこととなります。

さて、みなさんは、これから(もしくは、もうすでにお仕事等で)「データ、すなわち、数字を扱って、“問い”に対して適切な答えを導き出す(意思決定する)」機会に遭遇することとなります。具体例としては、

- ・どの科目を履修すべきか？

- ・ 期末テストを来週に控え、何を勉強すべきか？
- ・ 実施したイベントはいくらの効果があったのか？
- ・ 施設内で増加しつつある事故を最も低減可能な方策は何か？
- ・ 新商品の分配はどうするか？

など枚挙に暇がありません。みなさんはこれらの問に対して、どのようなプロセスで「データ」から「答え」を見出し、その「答え」を他者に伝えればよいかイメージが出来るでしょうか？このような能力は「問題解決力」の一部として、扱う「データ」量が増加の一途をたどる今日において重要なものと認識されており、その基礎を修得しておくことは自らの付加価値を生み出すものと考えます。

本科目は、受講なさる方々が数字（データ）を扱うことで、これから遭遇する新たな“問い”に対して「適切な答え」（意思決定）を見出し、それを他者に伝える「プロセス」を修得して頂くことを目的にしています。そのために、「データ分析」に用いられる語句や指標の意味や算出を正確に理解する（1単位目）、「データ分析」のプロセスに慣れる（2単位目）を設定しています。

なお、数字を扱うにあたって、数学に自信のない方もいらっしゃると思いますが、本科目ではデータ分析のプロセスを重視しますので、統計については必要最小限度（平均、中央値、最頻度、標準偏差、t検定程度）の知識しか用いません。しかも、それらの計算はExcelを用いることで解決しますので、ほぼ心配はいりません。（ただし、得られた計算結果の意味は理解していただく必要はあります）

1 単位め アドバイス

教科書や参考図書をよく読み、『客観式レポート集』記載の課題に解答してください。
「TFU オンデマンド」上で解答することも可能です。

2 単位め アドバイス

この課題は、データ分析の基本的なプロセスが身についているかを測るものです。
既に、「決めるべきこと」、および、「データ」は問題で与えられているので、あなたが決定すべきことは「評価軸」とその「要因」になります。そして、それを決めたら、欠損処理等のデータの前処理（クレンジング）を行ってデータ分析を行い、結果を解釈してください。その上で、それらから導かれたあなたの判断を説明してください。

なお、解答する際は教科書をよく読み込んで、それ記載されたプロセスを意識して行うようにしてください。

科目修了試験

■実施方法

会場試験は実施しません。大学からメールで送信された問題に、期限までに自宅で解答し、メールで提出する方法で実施します。

■申込・解答方法

- ① レポート課題（2課題）にすべて合格する（提出ではなく合格が条件です）。
- ② ur@tfu-mail.tfu.ac.jp（2単位めレポートの送信先アドレス）へ、下記事項を記載して科目修了

試験申込を行う（申込みは随時受付します）。

【件名】「データ分析とプレゼンテーション技法」科目修了試験申込

【本文】学籍番号・氏名・科目修了試験受験を希望する旨を記載

③申込みをしたメールアドレスに、大学から試験問題が送信される。

④期限までに解答し、上記③の試験問題送信メールに返信する形で提出（解答期限は試験問題送信後、約2週間）。

■評価基準

科目修了試験では、データ分析におけるプロセスを一通り問います。

与えられた問題に対して、以下の観点から評価します。

1. 分析の概念設定（作業仮説設定）に関する問いでは「何を決めるのか」「何で評価するのか」「評価に寄与する要因は何か」を①互いに矛盾なく決定できているか、②それぞれを設定した理由を他者に説明できているか、の観点から評価します
2. 分析に関する問いでは1.に基づいて、①収集したデータの適切さ、②分析方法の適切さ、③分析結果、の観点から評価します
なお、収集したデータについてはその出自を必ず記載してください
3. 分析結果の解釈・提示に関する問いでは、2.で得られた結果をそれぞれ適切に①図表化できているか、②解釈できているか、の観点から評価します
4. あなたの意思決定に関する問では、3.を基に決定したあなたの決定について①3.の解釈と整合しているか、②論理的な説明がなされているか、の観点から評価します

なお、データ分析とそれを基にした意思決定において、決定には絶対の「答え」はありません。あるのは、その決定に説得力があるか（適切なプロセスを踏んでいるか）なので、それを意識して解答してください。