

データ利活用 トレーニングブック

実践編

2024年度アシストソリューション研究会
東日本「ビジネスに貢献するデータ活用」分科会

目次

1. 概要.....	1
2. 問題の難易度の考え方.....	3
3. 業務時間分析 （難易度：★☆☆）.....	5
4. クラス成績向上 （難易度：★★☆）.....	32
5. 健康寿命延伸の検討 （難易度：★★★）.....	79

1. 概要

1-1. 本書の位置づけ

.....

本書は、シン・KKDによるデータ利活用を実践するための演習問題集である。

「03_データ利活用トレーニングブック_理論編.pdf」を参照しながら、実際に自身でも手を動かして本書の問題を解くことで、シン・KKDによるデータ利活用を実践できるようになってもらえれば幸いである。

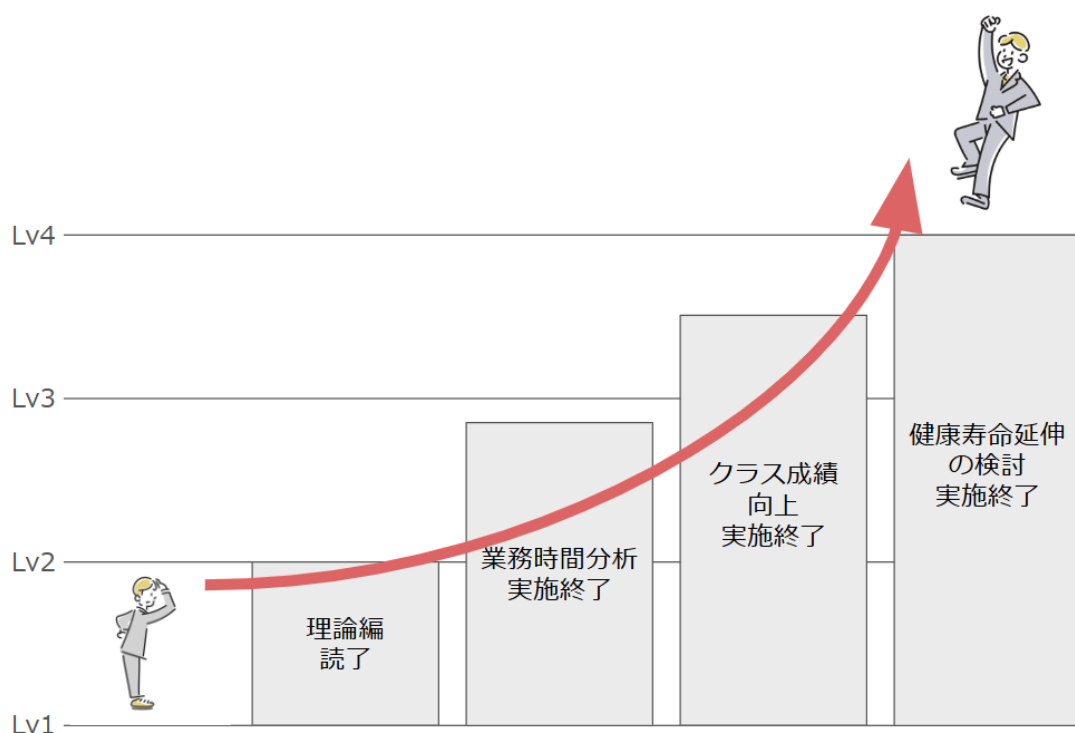
1-2. 問題一覧

下表に、本書に用意した演習問題を記載する。

難易度	タイトル	概要	実施目的
☆☆☆	業務時間分析	自身の業務内容・時間を分析し、スキルアップの時間を確保する方法を検討する	● 実践を通じてシン・KKDの流れを把握する
★★☆	クラス成績向上	テストの点数を分析し、点数を上げる方法を検討する	● あるべき姿の設定ができるようになる ● データ分類を意識した可視化およびグラフの特徴を踏まえた結論の整理ができるようになる
★★★	健康寿命延伸の検討	平均寿命・健康寿命の推移や健康の要素から、健康寿命を延ばす方法を検討する	● 論理的思考を用いて問題から課題を検討できるようになる ● 仮説検証に必要なデータを自分で考え、オープンデータを利用してデータ収集できるようになる

2. 問題の難易度の考え方

問題は、「02_データ利活用トレーニングブック_シン・KKDスキル一覧.xlsx」の必要スキルが段階的にレベルアップするように難易度を設定している。スキルアップのイメージを下図に示す。



「理論編」を読了している場合、すべてのスキルがレベル2になっている。

「実践編」の演習問題を実施していくことで、シン・KKDスキル一覧に記載しているスキルを向上させ、最終的にはすべてのレベルが4になることを目指して欲しい。

3. 業務時間分析

(難易度: ★★★)

3-1. 内容

.....

あなた自身がデータ利活用できるようになるためにスキルアップの時間を1週間のうち2割確保する必要がでてきた。

自分の現状の業務内容を分析して、スキルアップのための時間を1週間のうち2割(*)確保するためにはどうしたらいいかを検討する。

(*)今回は実践編の問題のため、2割という目標を設定している。2割の目標に近づくにはどうしたらよいかを考えよう。



3-2. 目標スキルレベル

.....

この演習問題は、実践を通じてシン・KKDの流れを把握することを目標としている。演習問題を解くことで、以下のスキルレベルに到達することを目指してほしい。

スキル	到達レベル
問題解決プロセス / 仮説設定プロセスの理解	★★★★☆
論理的思考能力	★★★★☆
データ構造の理解	★★★☆☆
データ収集力	★★★★☆
データ加工技術	★★★★☆
データ分析力	★★★★☆

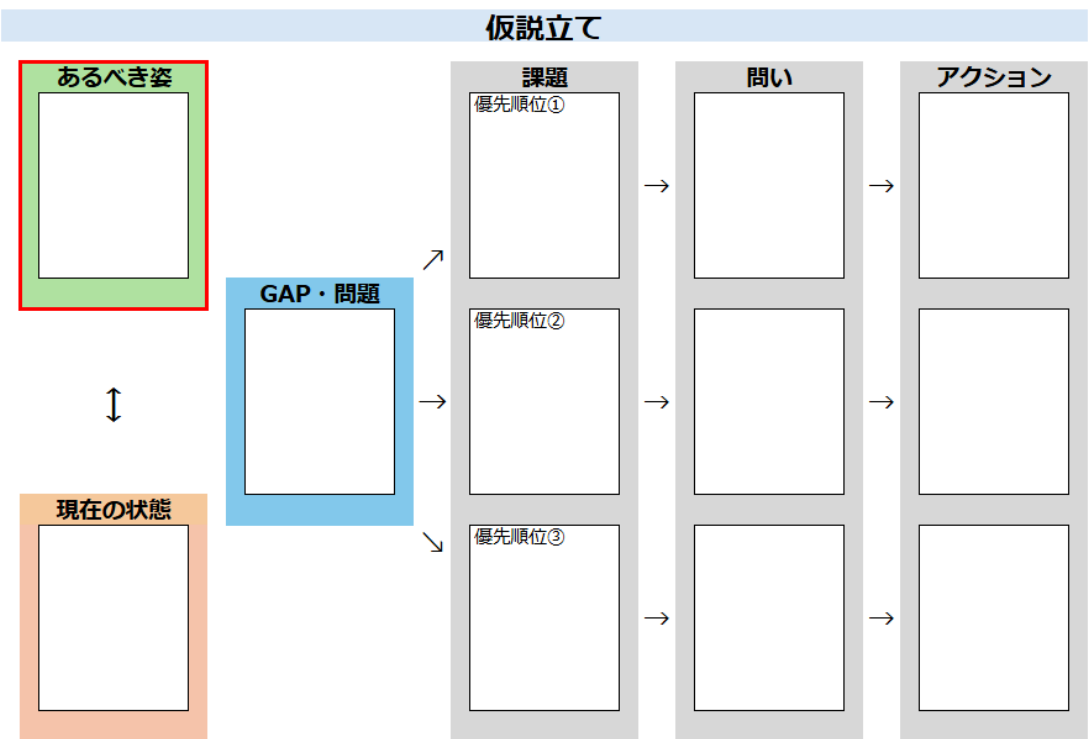
※詳細は「02_データ利活用トレーニングブック_シン・KKDスキル一覧.xlsx」を参照

3-3. 仮説立て

3-3-1. 理想(あるべき姿)の設定

はじめに、成し遂げたい目標を「あるべき姿」として設定する。何を「あるべき姿」とするかで、その後の「問題」や「課題」が変わってくるため、可能な限り定量的、定性的に具体的にありたい姿を記載するよう心がける。

○フレームワークに書いてみよう！



例)

あるべき姿

1週間のうち6時間をスキルアップ時間として確保

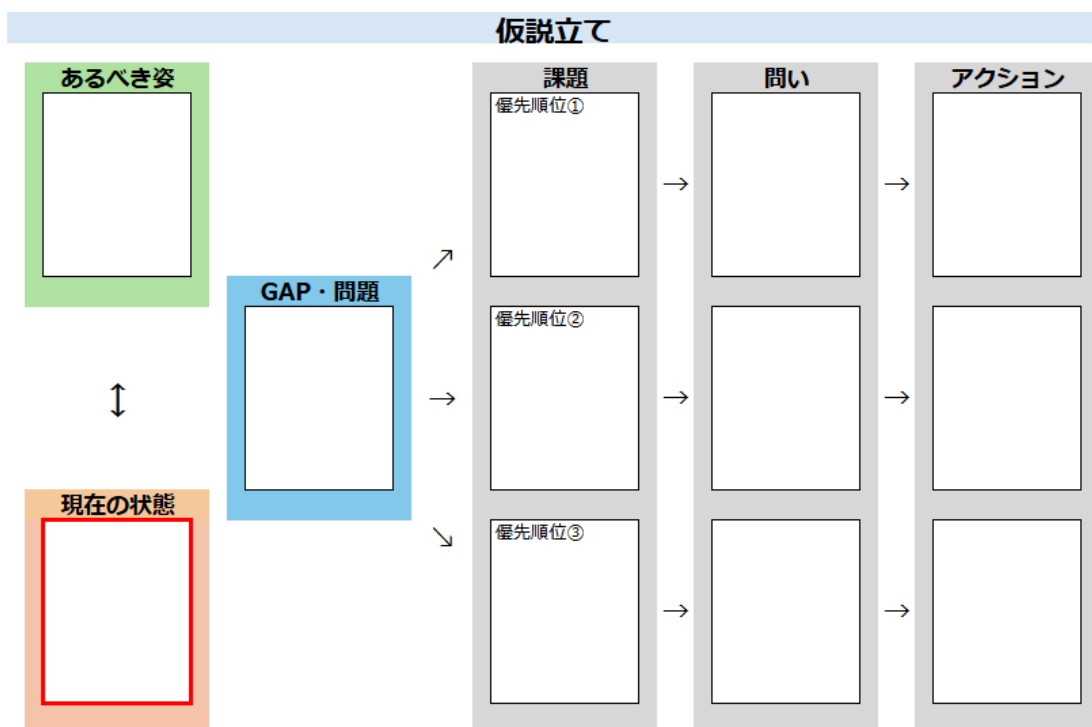
- ・いつ : 2024年9月～12月の間
- ・どこで : 業務時間内
- ・だれが : 自分
- ・何を : データ利活用に必要なスキル
- ・なぜ : 自己成長と所属チームへの貢献
- ・どのように : eラーニングシステムを利用

3-3-2. 現状整理

①現在の状態(現状)を記載

自身の業務やスキルアップについて、現在置かれている状況・事実を業務で使用している予定表やカレンダーから整理する。その際、対象の期間は1か月以上とする。対象期間が短い場合に、いつもそうなのか、その時はたまたまなのか、区別がつけにくいいためである。

○フレームワークに書いてみよう！



できるだけ多角的な視点から情報を記載する。必要に応じて関係者からも情報を収集する。

例)

現在の状態

<基本情報>

- ・勤務時間 6h/日
- ・勤務日数 20日/月
- ・テレワーク 10日/月
- ・通勤時間 1h/片道

<業務内容の内訳> ※2024年5～7月の実績

社内業務 184.17h

- ・MTG 86.75h (18.96%)
- ・イベント準備 4.50h (0.98%)
- ・プロジェクト活動 84.25h (18.42%)
- ・メーカー対応 6.00h (1.31%)
- ・その他 2.67h (0.58%)

顧客対応業務 145.17h

- ・アポ 107.33h (23.46%)
- ・社内顧客対応 37.83h (8.27%)

スキルアップ 69.67h

- ・ソ研 49.58h (10.84%)
- ・社内勉強会 20.08h (4.39%)

休暇 58.50h (12.79%)

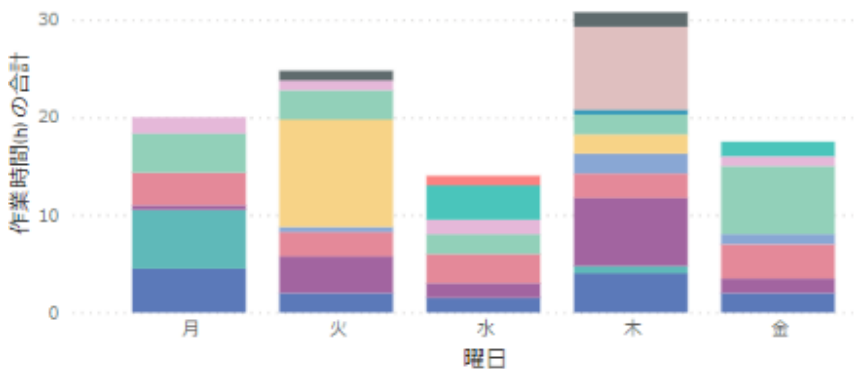
②現状に対するデータ分析を実施

手元にデータが既にある場合、現状をより詳細に把握するために、可視化を行う。何にどれくらいの時間を費やしているのかがわかるようにグラフ化すると現状の把握が容易になる。

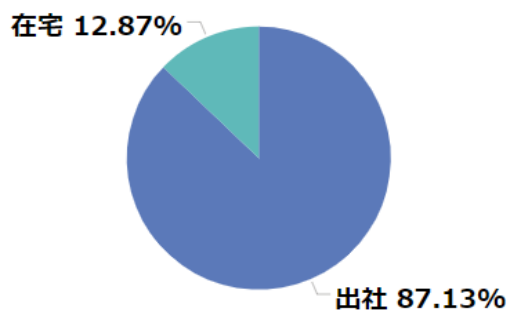
例)

4月の曜日ごとの業務スケジュール内訳

カテゴリ ●朝会 ●グループ内... ●社内打合... ●社内打合... ●社外打合せ ●外部セミ... ●個人作業



4月の出社/在宅の割合



数値化やグラフ化にあたり、予定表やカレンダーの内容を要素ごとに分類しデータとして準備をすると、要素ごとの集計や割合の算出などが容易になる。

例)

2024/05/01(水) 9:00～10:00 チームの定例MTG

↓

年月日	曜日	開始時間	終了時間	所要時間	内容
2024/05/01	水曜日	9:00	10:00	1h	社内定例MTG

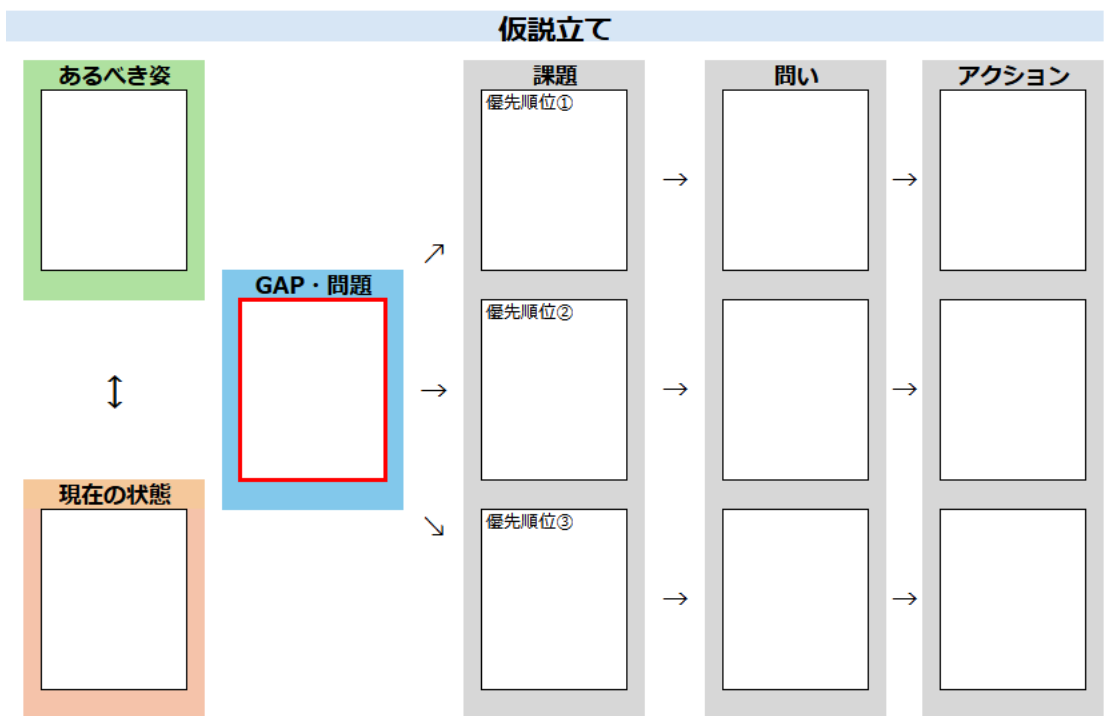
なお、数値化・グラフ化する際は、以下のようなツールを使うとデータの加工が便利のため、利用について適宜検討する。

- ・BIツール
- ・Excel、スプレッドシート など

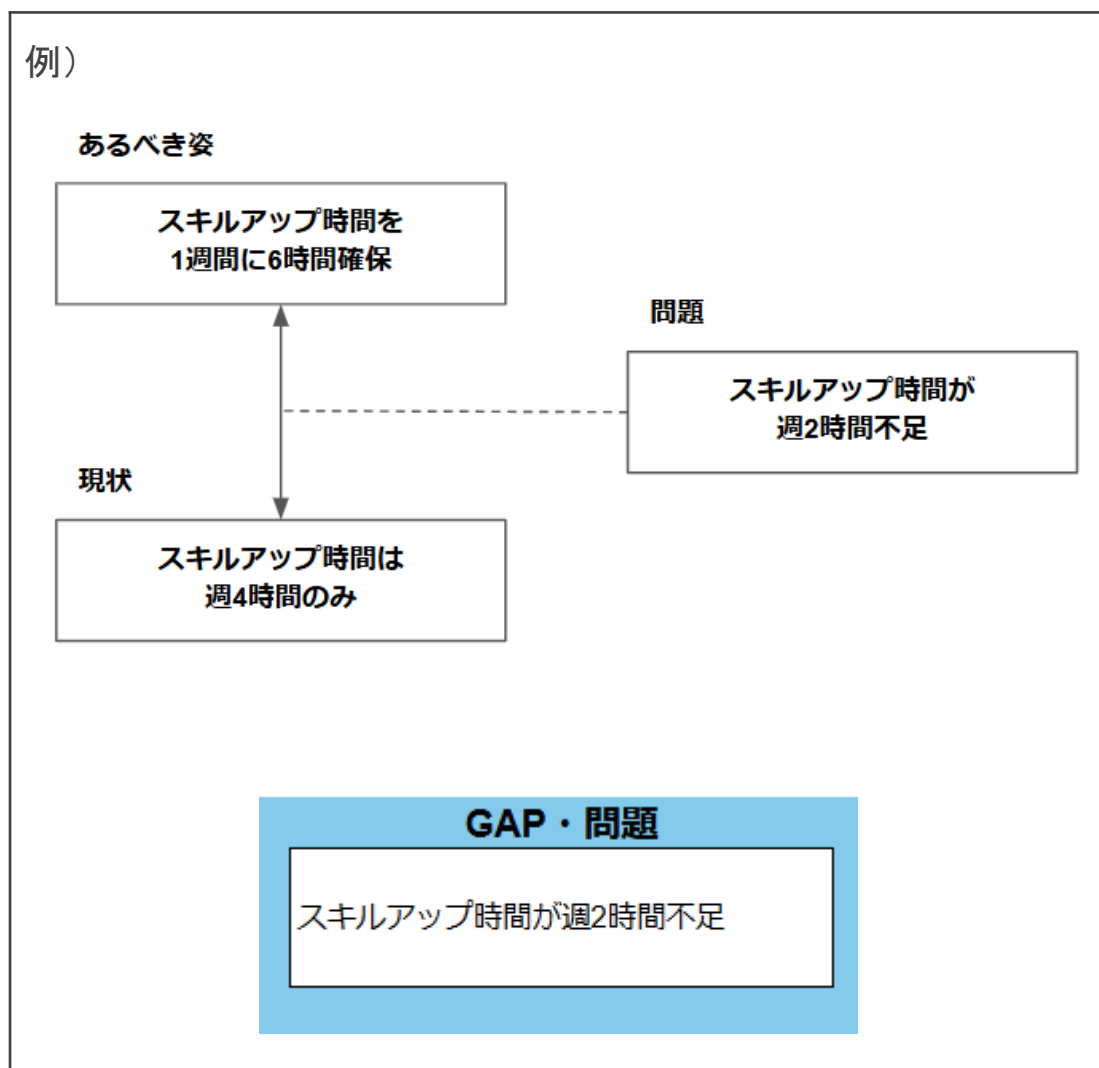
3-3-3. GAP・問題の設定

設定した「あるべき姿」と整理した「現状」の差分（「コト」や「状況」）を「GAP・問題」として書き出す。

○フレームワークに書いてみよう！



ここでは、シンプルに「あるべき姿」と「現状」の差を書き出す。曖昧な表現、憶測や主観的な表現は避ける。

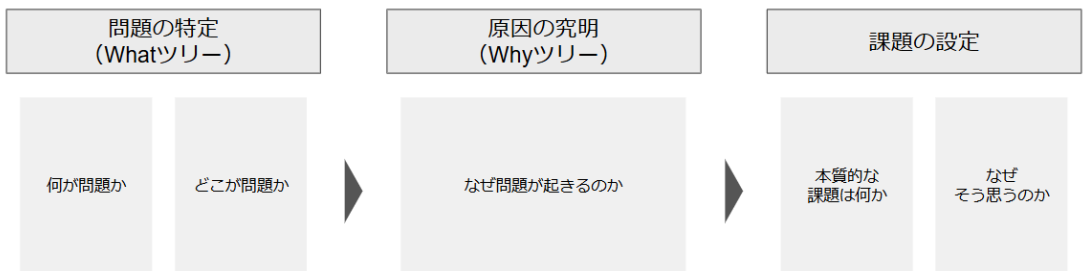


3-3-4. 課題の設定

導き出された「問題」が、どこで起きているのかを明らかにし、「課題」に落とし込む。

その際、ロジックツリーを用いて、物事の要素を分解して網羅的に洗い出すことで問題の発生個所を特定していく。MECEを意識して、問題の全体像や関係性を細かく分解していく。

○ロジックツリー記載用シートに書いてみよう！



記載する際は、「問題の特定」「原因の究明」「課題の設定」を意識するとよい。

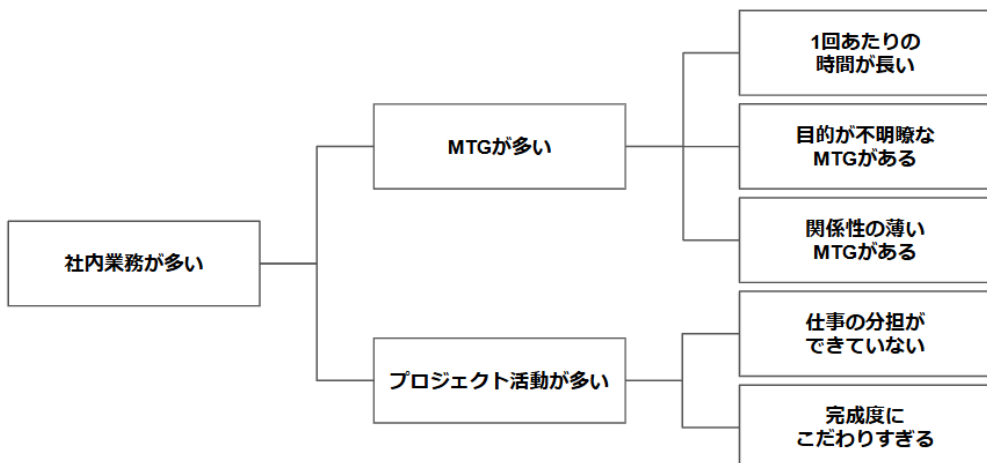
①ロジックツリーの作成

ロジックツリーを作成するときは、設定したGAP・問題から原因を究明する。原因を引き起こしている要因を3つ考えるとヌケモレを減らしやすい。原因究明が浅くならないよう、ロジックツリーを活用しながら「なぜ？」を繰り返し深堀する。

例) Whatツリー(スキルアップ時間不足の要因は?)



例) Whyツリー(なぜ社内業務が多いのか?)



あるいは、実際に自分の予定と理想の予定を書き出してみるのもおすすめです。どこに問題があるのか視覚的に把握できるため、比較することで問題の所在がわかりやすくなる。

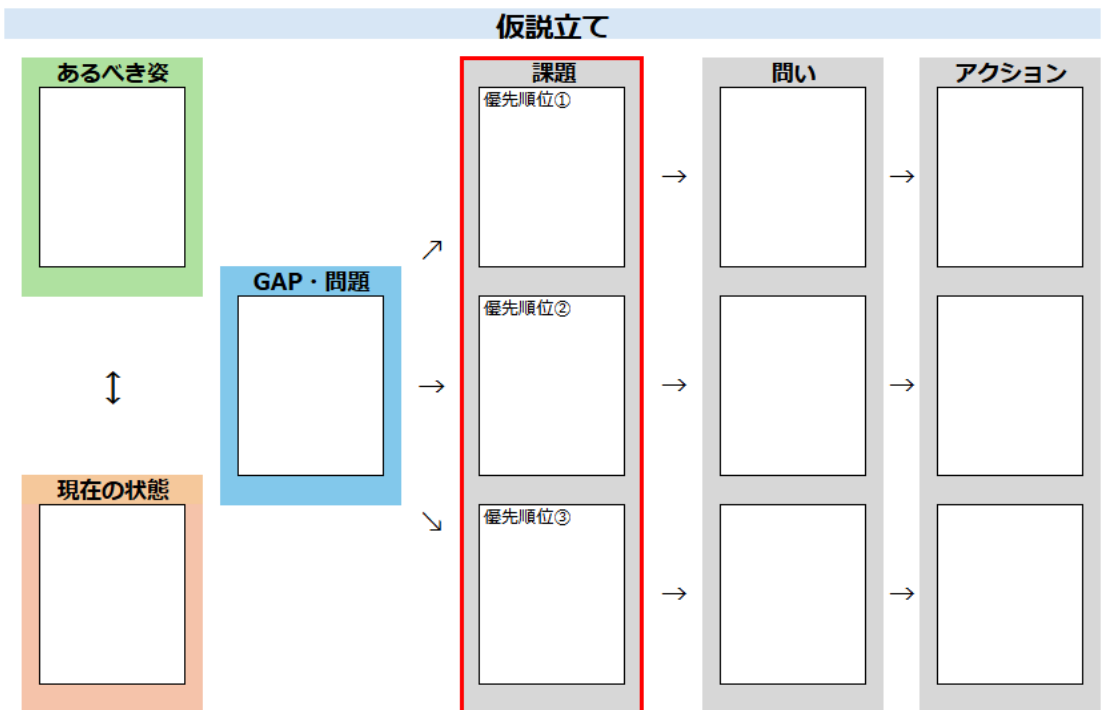
例)

		7月1日 月		7月2日 火	
		理想	現実	理想	現実
出社 or 在宅		出社	出社	在宅	在宅
開始	終了				
8:00	8:30	データ利活用			
8:30	9:00	朝会	朝会	予備	朝会
9:00	9:30	プロA 作業		部内打合せ	部内打合せ
9:30	10:00		プロA 作業		
10:00	10:30	社内ユーザ打合せ	社内ユーザ打合せ	プロA 作業	プロA 作業
10:30	11:00	予備	プロA 作業続き		
11:00	11:30	G内進捗MTG	G内進捗MTG	社内ユーザ打合せ	社内ユーザ打合せ
11:30	12:00				
12:00	12:30	お昼休憩	お昼休憩	お昼休憩	お昼休憩
12:30	13:00				
13:00	13:30	プロA関連	プロA関連		
13:30	14:00	ベンダー打合せ	ベンダー打合せ	プロA 作業	プロB 作業
14:00	14:30				
14:30	15:00	社内ユーザ打合せ	社内ユーザ打合せ		
15:00	15:30		プロB 作業	データ利活用	問い合わせ対応
15:30	16:00	プロB 作業	問い合わせ対応		
16:00	16:30			個別作業	個別作業
16:30	17:00	社内ユーザ打合せ	社内ユーザ打合せ		
17:00	17:30	データ利活用	プロB 作業続き		
17:30	18:00				

②課題を記載

原因究明のステップで作成したロジックツリーの内容から、本質的な課題は何なのか抽出していく。いくつかの要因が関係している場合もあるため、判断に迷う場合は、原因究明のステップで作成したロジックツリーの内容を切り口に、予定表やカレンダーを改めて確認することもおすすめである。

○フレームワークに書いてみよう！



「問題を解決するために実行すべきこと」を課題として設定する。
その際、自分の対応可能な範囲に限定せず、外部の事例や他者の
アイデアを取り入れると良い。

例)

課題

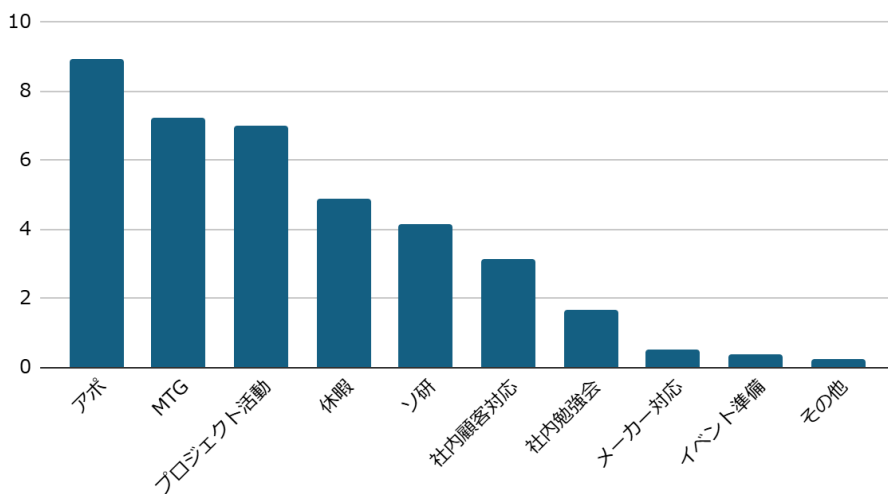
優先順位①

無駄なMTGをなくす、もしくは
MTGの無駄を減らす

さらに、データ可視化の結果から読み取れる課題であるか(それが課題であるエビデンスがあるか)を確認する。

例)

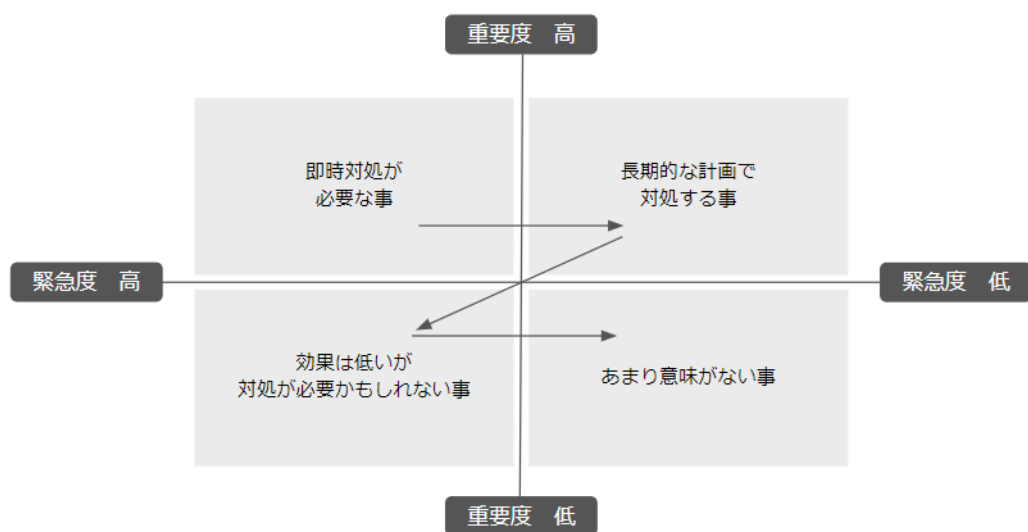
業務種別 (1週間当たりの平均)



※社内業務のうちMTGに割いている時間は多い
(平均すると1週間あたり7時間程度MTGの時間がある)

また、課題が複数ある場合は、問題への影響度(価値)や、必要コスト、短期・中期的な視点から優先順位を決めることが望ましい。

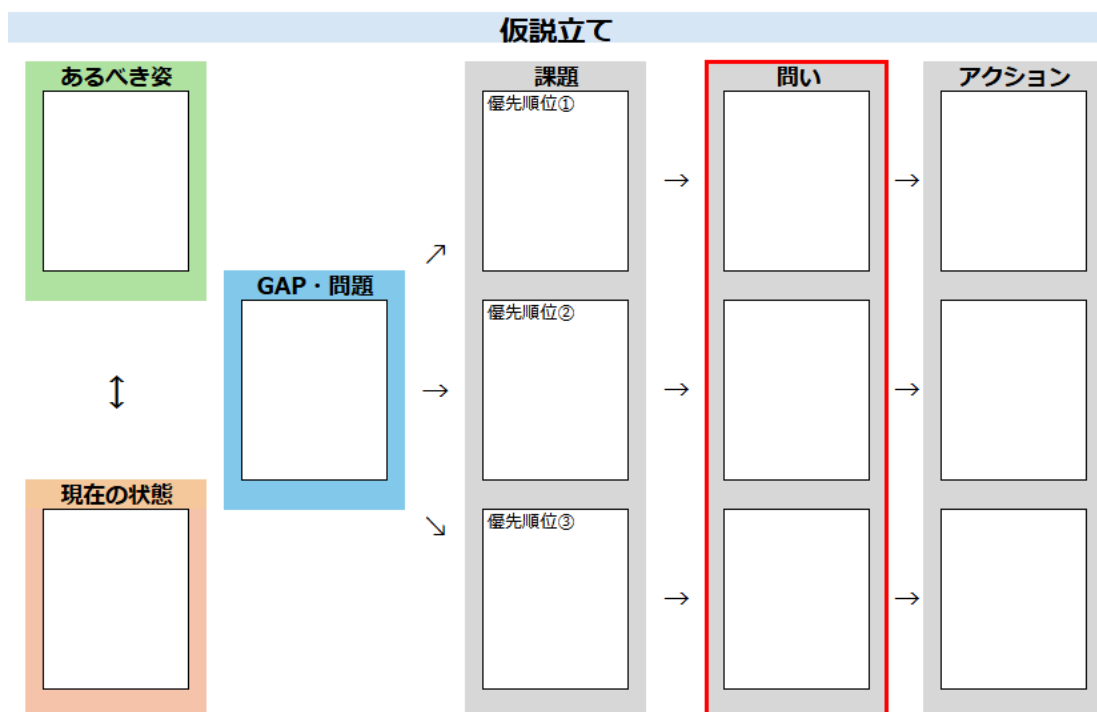
例)



3-3-5. 問いの設定

「あるべき姿」と現状とのギャップを考えながら「問い」を考える。

○フレームワークに書いてみよう！



何を改善すれば良いか、具体的な行動に移せるか、定量的・定性的に整理されているかを意識する。

例)

課題

優先順位①
無駄なMTGをなくす、もしくは
MTGの無駄を減らす



問い

参加が必須でないMTGもあるのでは
ないか

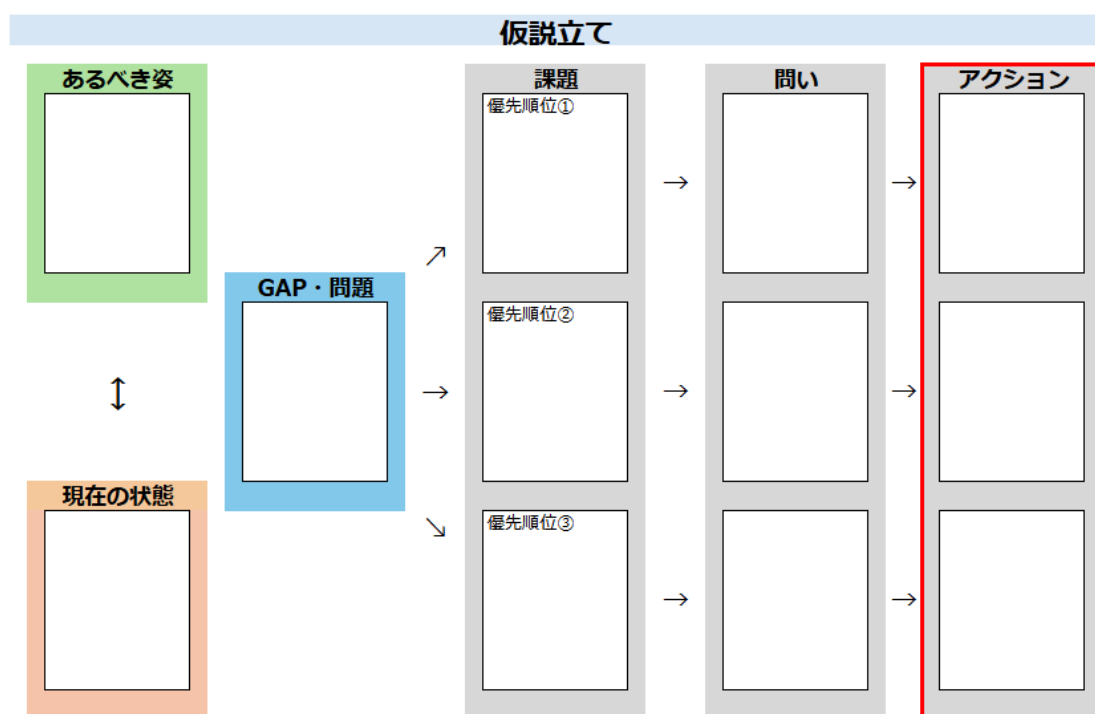
情報共有や集まることが目的になって
いないか

時間を短縮できる余地があるのでは
ないか

3-3-6. アクションの設定

「3-3-5. 問いの設定」で記載した、課題を解決するための「アクション」を記載する。

○フレームワークに書いてみよう！



「アクション」の記載にあたり、「誰が、いつまでに」が明確になるように記載を行う。

「アクション」を踏まえて、現状からあるべき姿に近づくことができるのかを考え、明日からすぐ始められるような具体的な内容・中長期的な観点で実施が必要なアクションなのかを整理することがポイントである。アクション欄に「すぐできる」、「効果が出るのに時間がかかる」等の観点で整理すると良い。

例)

問い

参加が必須でないMTGもあるのではないか
情報共有や集まることが目的になっていないか
時間を短縮できる余地があるのではないか



アクション

目的をはっきりさせたいうえで参加するMTGを決める
情報共有はチャットやプロジェクト管理ツールで行う
事前にアジェンダを作成し、議事内容や終了条件を共有する
※すぐできるが、メンバーの協力も必要

3-4. 仮説検証の計画

3-4-1. 検証方法の設定

仮説検証の計画では、「アクション」によって課題が解決されるのかを確かめるための方法を検討していく。

3-4-2. データ収集方法の設定

本問題については、「3-3-2. 現状整理」で実施した、数値化やグラフ化の中で、各自の予定表やカレンダーの内容を要素ごとに分類し、データとして準備した。

3-4-3. データ可視化、得たい結論の設定

今回は、「3-3-6. アクションの設定」で設定したアクションを実施したのち、「3-3-2. 現状整理」と同じデータを用いて、1週間のうち2割の時間を確保できたのかを検証していく。

そのため、本問題においては「仮説検証の計画」の実施は不要とする。

3-5. データ収集

3-5-1. データの取得

3-5-2. データの加工

本問題については、「3-3-2. 現状整理」で実施した、数値化やグラフ化と同様に、各自の予定表やカレンダーの内容を要素ごとに分類してもらい、データを収集・可視化する。

そのため、本問題ではデータ収集の手順を省略する。

3-6. データ可視化

3-6-1. データの可視化

3-6-2. データの分析

アクションを実行した結果、1週間のうち2割の時間を確保することができたのかを検証する。そのため「3-3-2. 現状整理」と同様に、業務時間のデータを再度収集し、データの可視化・分析を行っていく。

3-7. 結論・考察

3-7-1. 可視化結果を記載

① 可視化結果を記載

「3-6. データ可視化」から読み取った結論をフレームワークに記載する。

② 仮説との突合

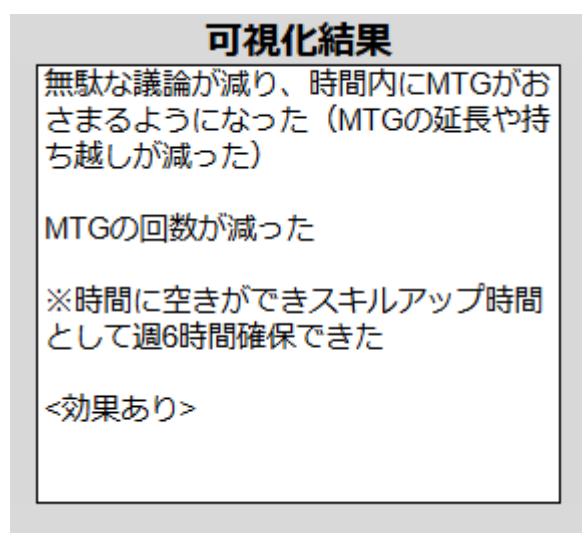
「3-3-6. アクションの設定」で実施したアクション結果の振り返りをもとに、自身で設定した仮説に対してのアクションが正しかったのか・次のアクションが必要なのかをフレームワークに記載する。

○フレームワークに書いてみよう！

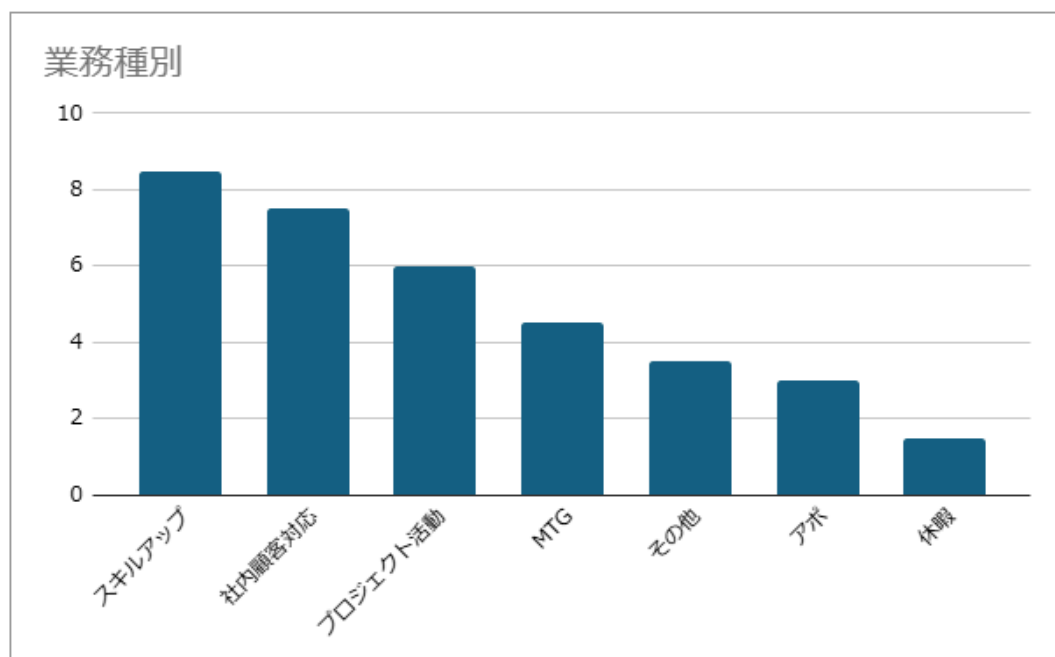
		仮説検証の計画		結論・考察	
問い	アクション	計画		可視化結果	
	→	<ul style="list-style-type: none">・検証方法・検証データ・データ収集方法・可視化内容・可視化した結果得たい結論	→		
	→	<ul style="list-style-type: none">・検証方法・検証データ・データ収集方法・可視化内容・可視化した結果得たい結論	→		
	→	<ul style="list-style-type: none">・検証方法・検証データ・データ収集方法・可視化内容・可視化した結果得たい結論	→		

1週間のうち2割確保できたのかを前提に、自身で設定したアクション結果を振り返ろう。できなかった場合は、なぜできなかったのかの原因を深ぼっていくことが大事である。

例)



検証（2024年9月の1週間）



検証開始からまだ1週間しか経過していないため単純比較はできないが、スキルアップの割合時間が増え、MTG時間が減っていることがわかる

「効果があった(→仮説が正しかった)」場合は、終了となる。「効果がなかった(→仮説が正しくなかった)」場合は、ネクストアクションを実施する。

ネクストアクションを実施するには、「仮説立て」から「結論・考察」までのプロセスを繰り返し実践することが必要である。

以上で「業務時間分析」は終了とする。



お疲れ様でした！！

4. クラス成績向上

(難易度: ★★★)

4-1. 内容

.....

あなたは、とある生徒が50人いるクラスAの担任である。

1学期、2学期の10科目のテストの点数結果(*)が手元にある。

このデータをもとに3学期目の10科目の全体のテスト点数の平均値を上げるにはどうしたら良いか？

(*)クラスAの1学期・2学期のテストの点数結果は、添付のZIPファイルに同封している。ZIPファイルを解凍し、[実践編_クラス成績向上.csv]をダウンロードのうえ問題に取り組むこと。



4-2. 目標スキルレベル

この演習問題は、以下の状態になることを目的としている。

- ・あるべき姿の設定ができるようになる
- ・データ分類を意識した分析およびグラフの特徴を踏まえた結論の整理ができるようになる

演習問題を解くことで、以下のスキルレベルに到達することを目指してほしい。

スキル	到達レベル
問題解決プロセス / 仮説設定プロセスの理解	★★★★
論理的思考能力	★★★★☆
データ構造の理解	★★★★☆
データ収集力	★★★★☆
データ加工技術	★★★★
データ分析力	★★★★

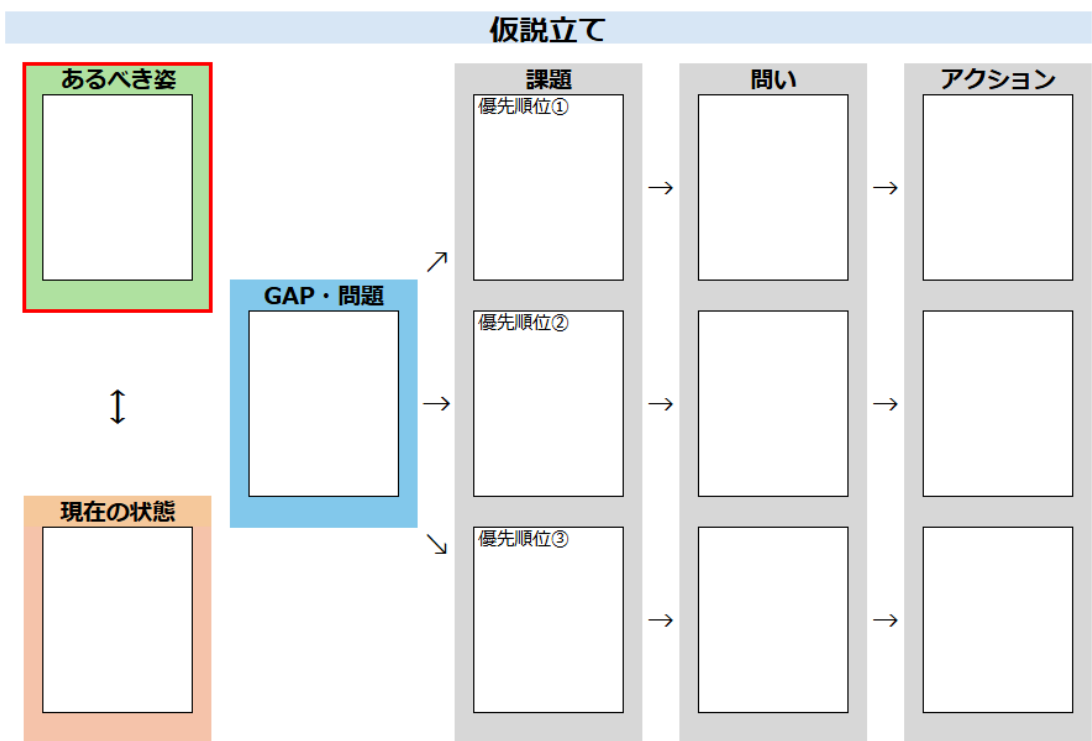
※詳細は「02_データ利活用トレーニングブック_シン・KKDスキル一覧.xlsx」を参照

4-3. 仮説立て

4-3-1. 理想(あるべき姿)の設定

はじめに、成し遂げたい目標を「あるべき姿」として設定する。

○フレームワークに書いてみよう！



具体的な目標値がない場合は、可能な限り思い浮かぶ理想を仮で設定する。判断に迷う場合は、関係者から意見を収集し、認識にずれがないか確認することも肝要である。

例)

あるべき姿

50人全員が3学期、10科目すべて100点を取る

4-3-2. 現状整理

①現在の状態(現状)を記載

生徒のテスト結果について、現状分かっていることを整理する。

○フレームワークに書いてみよう！

仮説立て				
あるべき姿		課題	問い	アクション
<div>あるべき姿</div>	GAP・問題	優先順位①	→	→
		優先順位②	→	→
		優先順位③	→	→
現在の状態				

例)

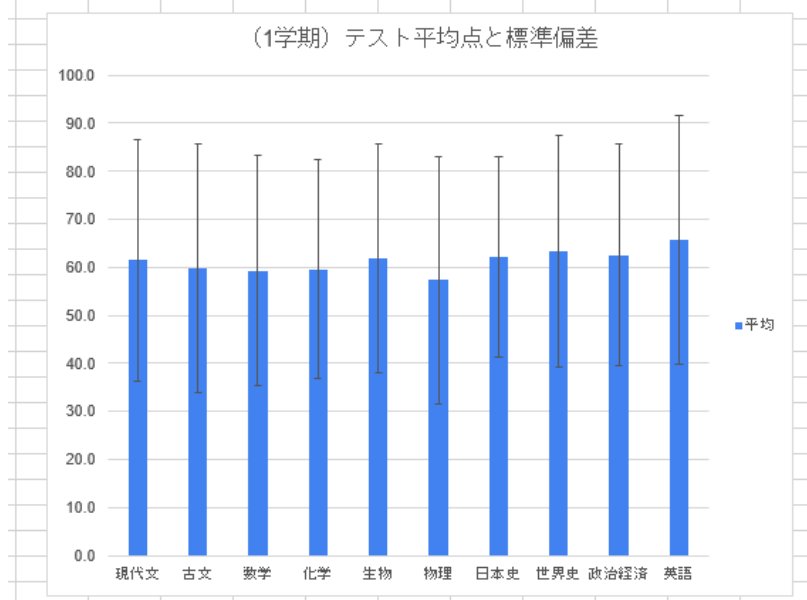
- ・各科目の平均点と標準偏差
- ・各科目の点数のばらつき
- ・各科目に対するテスト結果の相関関係

整理したい内容についてグラフ化すると、視覚的に現状について把握できる。

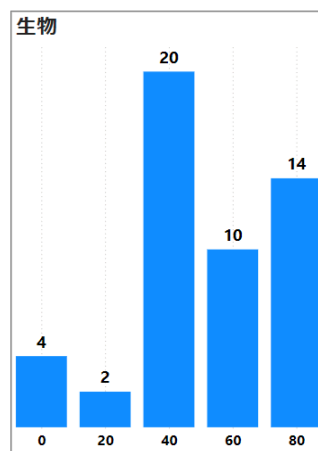
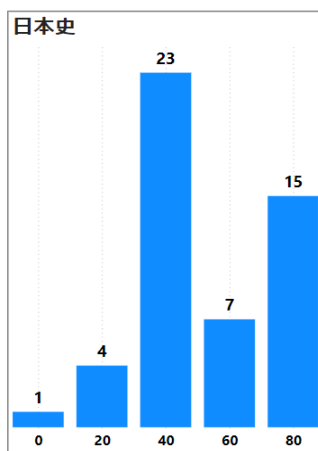
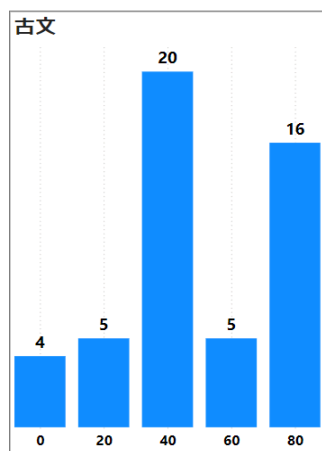
例) ※一部抜粋

・各科目の平均点と標準偏差【図1】

	現代文	古文	数学	化学	生物	物理	日本史	世界史	政治経済	英語
平均	61.5	59.8	59.3	59.5	62.0	57.3	62.2	63.5	62.6	65.7
標準偏差	25.1	25.9	24.0	22.8	23.8	25.7	20.9	24.1	23.0	25.9



・各科目の点数のばらつき【図2】

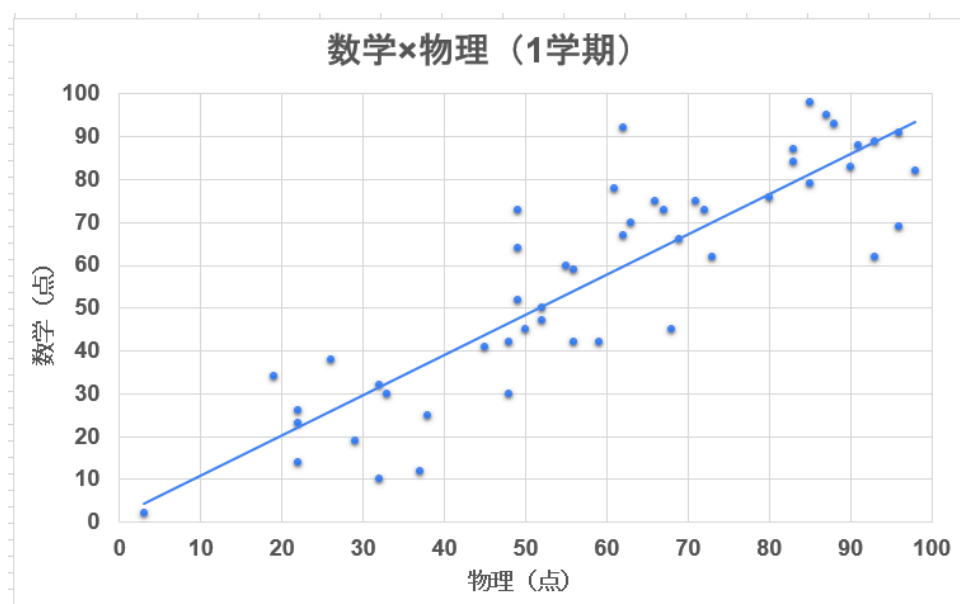


・テスト結果の相関関係(ヒートマップ図)【図3】

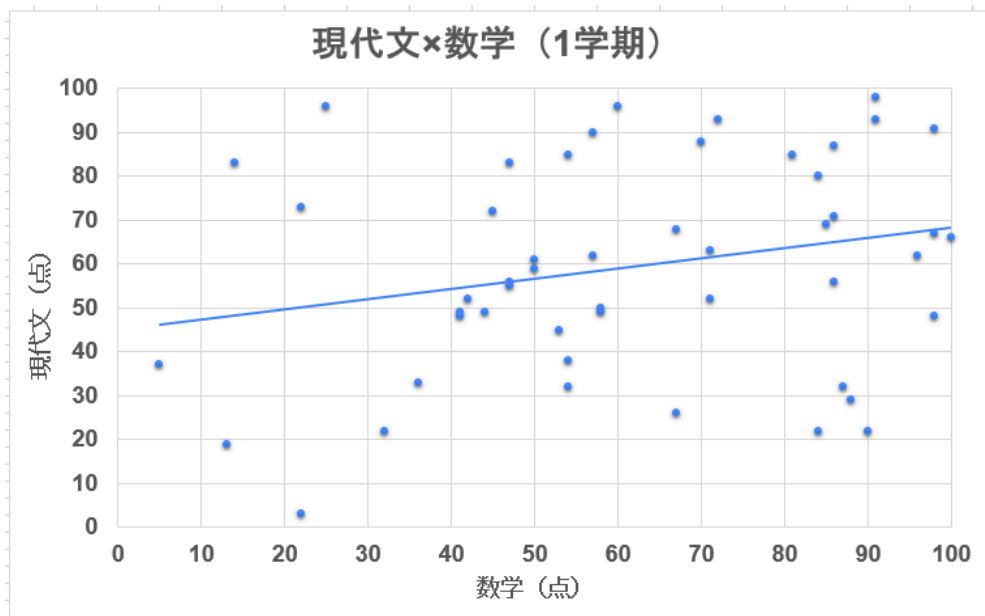
	現代文	古文	数学	化学	生物	物理	日本史	世界史	政治経済	英語
現代文	1.00	0.74	0.24	0.40	0.55	0.23	0.57	0.52	0.51	0.70
古文	0.74	1.00	0.26	0.43	0.49	0.35	0.64	0.57	0.63	0.70
数学	0.24	0.26	1.00	0.73	0.75	0.88	0.10	0.04	0.02	0.16
化学	0.40	0.43	0.73	1.00	0.75	0.78	0.23	0.25	0.26	0.30
生物	0.55	0.49	0.75	0.75	1.00	0.77	0.39	0.32	0.40	0.48
物理	0.23	0.35	0.88	0.78	0.77	1.00	0.18	0.18	0.15	0.21
日本史	0.57	0.64	0.10	0.23	0.39	0.18	1.00	0.85	0.84	0.53
世界史	0.52	0.57	0.04	0.25	0.32	0.18	0.85	1.00	0.72	0.49
政治経済	0.51	0.63	0.02	0.26	0.40	0.15	0.84	0.72	1.00	0.47
英語	0.70	0.70	0.16	0.30	0.48	0.21	0.53	0.49	0.47	1.00

・テスト結果の相関関係(散布図)【図4】

＜相関性のある散布図(文系と文系、理系と理系)＞



< 相関性のない散布図(文系と理系) >



例)

現在の状態

50名の1学期、2学期の10科目のテストの点数がある。

10科目：現代文、古文、数学、化学、生物、物理、日本史、世界史、政治経済、英語

※zipファイル [test_score.csv] 参照

10科目すべて、1学期も2学期も平均点は60点前後

※【図1】参照

10科目すべて、1学期も2学期も平均60点前後から±20~25点のズレがある

※【図1】参照

全体的にどの科目も「成績上位」「平均並み」「成績下位」の3グループに分けられる

※【図2】参照

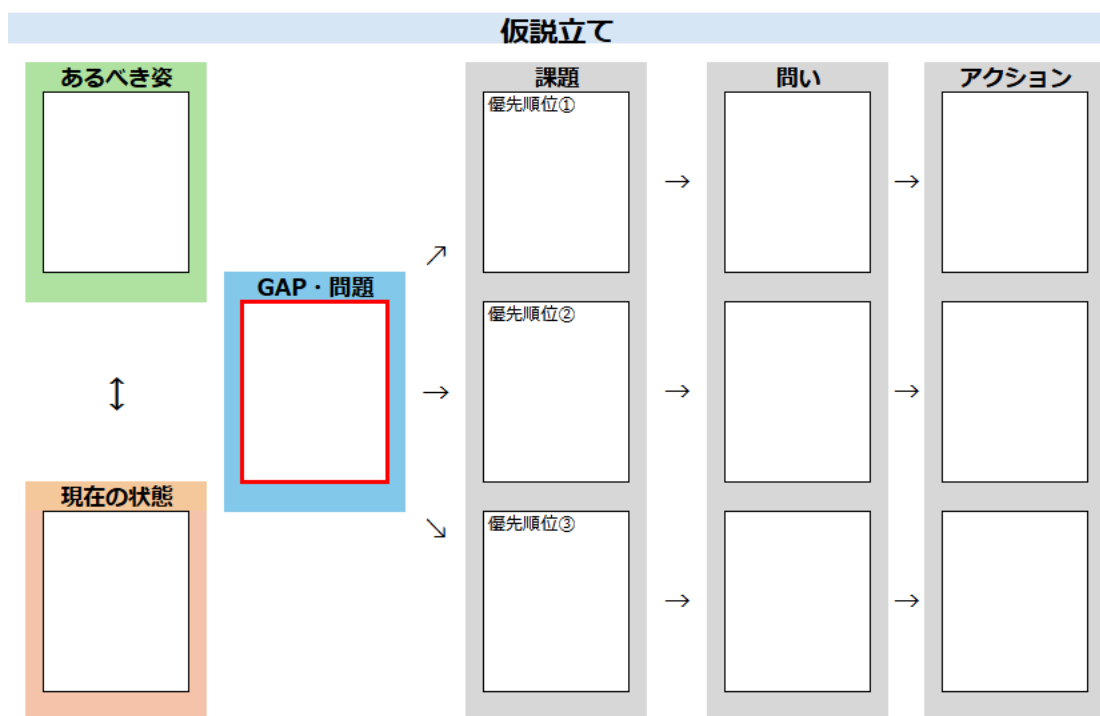
文系科目と理系科目で、点数に相関がある。これにより、クラスを「どの科目も平均並み」「文系科目が得意」「理系科目が得意」の3つに分類できる可能性がある。

※【図3】 【図4】 参照

4-3-3. GAP・問題の設定

設定した「あるべき姿」と、整理した「現在の状態」の差分（「コト」や「状況」）を「GAP・問題」として書き出す。

○フレームワークに書いてみよう！



例)

GAP・問題

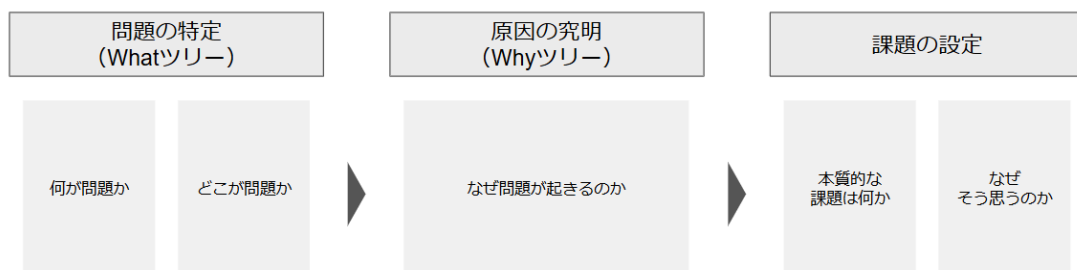
10科目すべて、平均点と比較すると約40点の差がある。

4-3-4. 課題の設定

導き出された「問題」が、どこで起きているのかを明らかにし、「課題」に落とし込む。

その際、ロジックツリーを用いて、物事の要素を分解して網羅的に洗い出すことで問題の発生個所を特定していく。MECEを意識して、どこが問題になっているかを細かく分解していく。

○ロジックツリー記載用シートに書いてみよう！

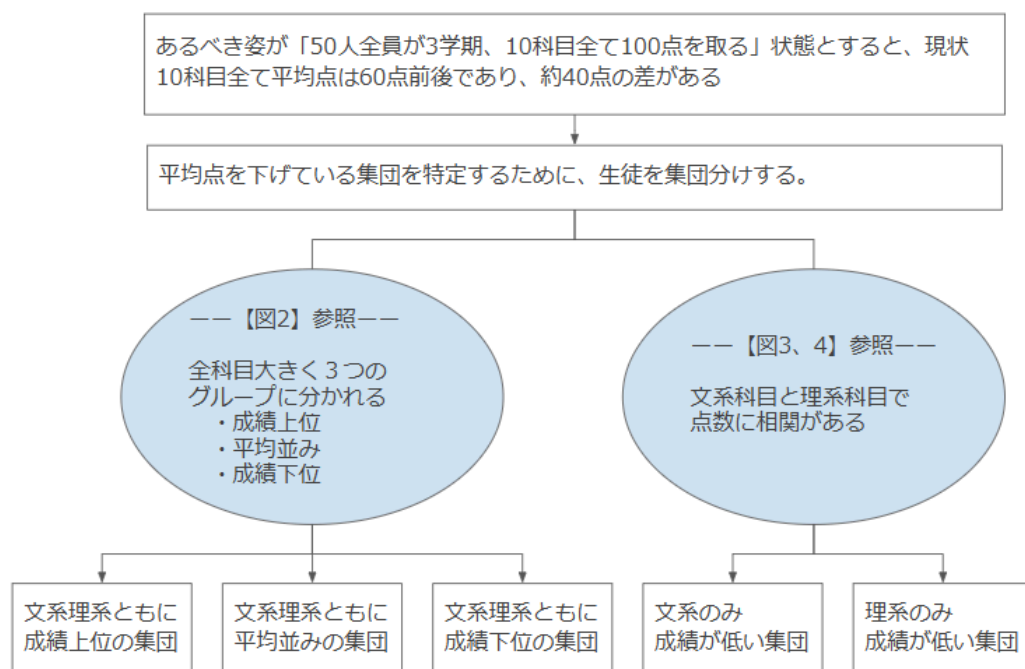


記載する際は、「問題の特定」「原因の究明」「課題の設定」を意識するとよい。

①ロジックツリーの作成

ロジックツリーを作成するときは、設定したGAP・問題から原因を究明する。

例) Whatツリー(平均点を下げている要因は?)

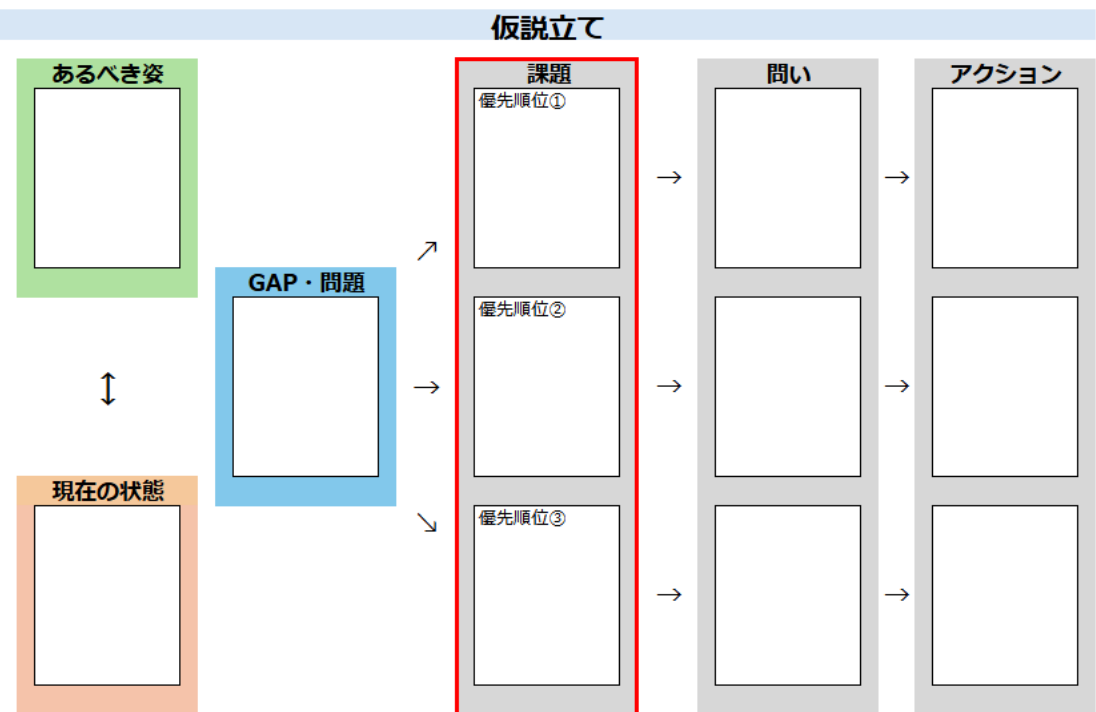


※今回、テスト結果以外のデータがないため、平均点を下げている原因(各生徒の学習時間など)を追及するのは不可能である。そのため「Whyツリー」は作成しない。

②課題の記載

原因究明のステップで作成したロジックツリーの内容から、本質的な課題は何なのか抽出していく。

○フレームワークに書いてみよう！



例)

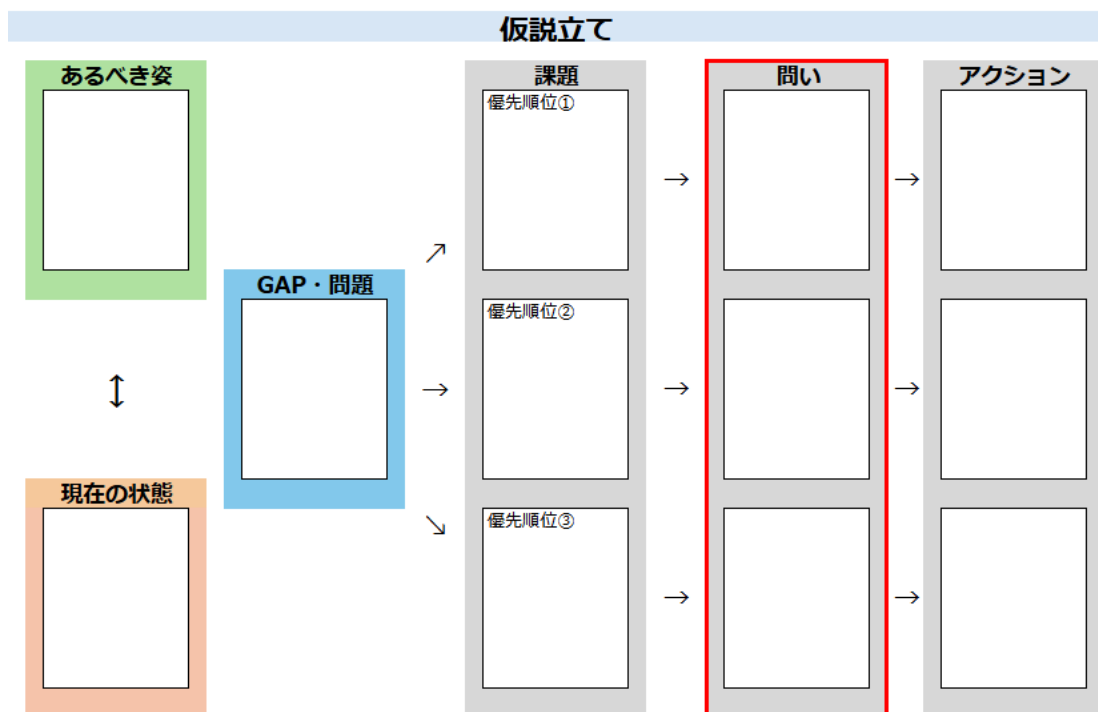
課題	
優先順位①	全体の平均点を下げている成績下位集団の点数を優先的に上げる必要がある
優先順位②	文系 / 理系科目それぞれ苦手な集団に分けて、適切な補講を行う必要がある

なお、データ可視化の結果から読み取れる課題であるか(それが課題であるエビデンスがあるか)を確認する。

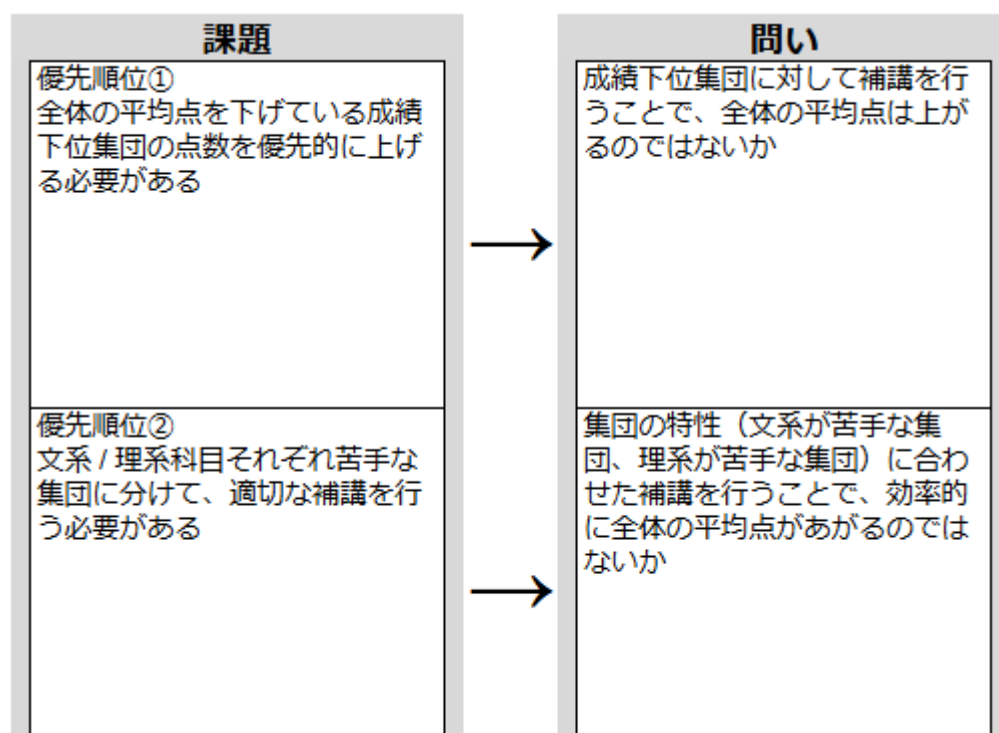
4-3-5. 問いの設定

「あるべき姿」と現状とのギャップを考えながら「問い」を考える。

○フレームワークに書いてみよう！



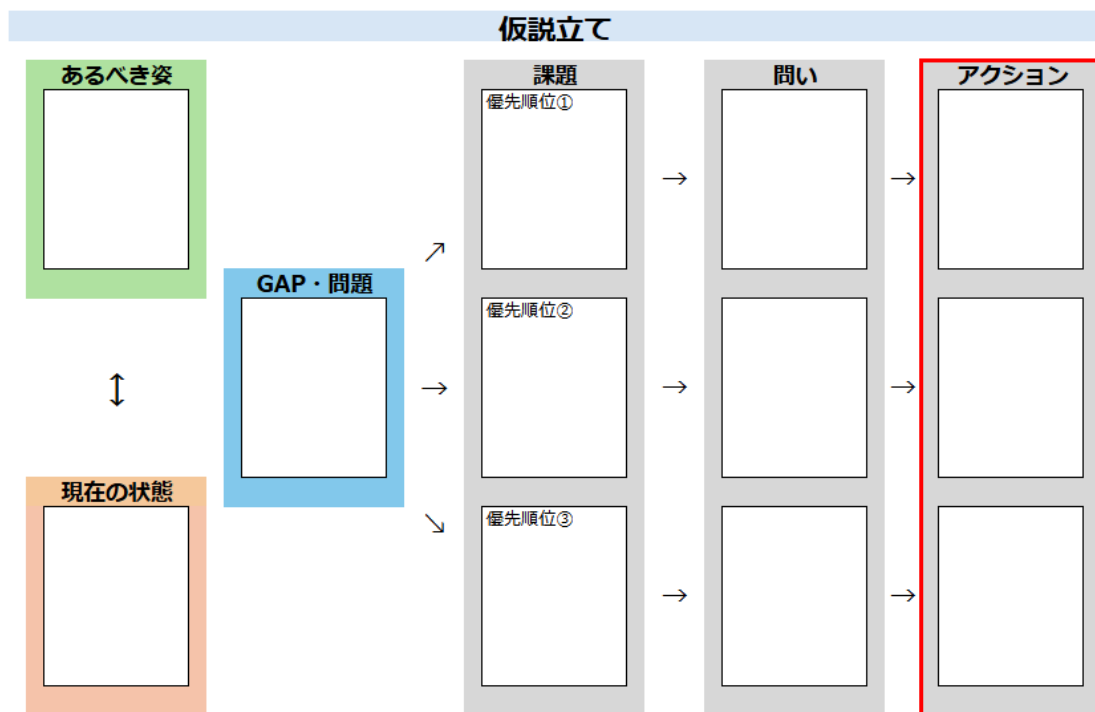
例)



4-3-6. アクションの設定

「4-3-5. 問いの設定」で記載した、課題を解決するための「アクション」を記載する。「アクション」の記載にあたり、「誰が、いつまでに」が明確になるように記載を行う。

○フレームワークに書いてみよう！



例)

アクション

効率的に補講を実施するために、成績のレベルにおよび特性（文系、理系）に準じたグループ分けを行う

4-4. 仮説検証の計画

4-4-1. 検証方法の設定

「アクション」により課題が解決するのか、を検証する方法を検討し、その内容を記載する。

○フレームワークに書いてみよう！

		仮説検証の計画	結論・考察
問い	アクション	計画	可視化結果
	→	<ul style="list-style-type: none">・検証方法・検証データ・データ収集方法・可視化内容・可視化した結果得たい結論	→
	→	<ul style="list-style-type: none">・検証方法・検証データ・データ収集方法・可視化内容・可視化した結果得たい結論	→
	→	<ul style="list-style-type: none">・検証方法・検証データ・データ収集方法・可視化内容・可視化した結果得たい結論	→

■検証方法

本問題は実践編の問題であり、3学期のテストの点数を把握することはできない。そのため、手元に与えられている「1学期と2学期のテストの点数結果」を使用して、特性（文系、理系）および成績のレベルに準じたグループ分けを実施する。

例)

計画

・検証方法

手元に与えられている「1学期と2学期のテストの点数結果」を使用して、特性（文系、理系）および成績のレベルに準じたグループ分けを実施する

4-4-2. データ収集の方法の設定

① 検証データを記載

検証に必要な「データ」を検討し、記載する。

■ 検証データ

本問題では3学期のテストの点数を把握することはできないため、手元に与えられている「1学期と2学期のテストの点数結果」を使用して検証を実施する。

例)

計画

- ・ 検証方法

手元に与えられている「1学期と2学期のテストの点数結果」を使用して、特性（文系、理系）および成績のレベルに準じたグループ分けを実施する

- ・ 検証データ

「1学期と2学期のテストの点数結果」を使用して検証を実施

②データ収集方法を記載

検証に必要な「データ」の収集方法を検討し、記載する。

主なデータ収集方法については、「03_データ利活用トレーニングブック_理論編」の「4. データ収集」を参照してほしい。

■データ収集方法

本問題では新たなデータ収集は不要のためスキップする。

例)

計画

- ・ 検証方法

手元に与えられている「1学期と2学期のテストの点数結果」を使用して、特性（文系、理系）および成績のレベルに準じたグループ分けを実施する

- ・ 検証データ

「1学期と2学期のテストの点数結果」を使用して検証を実施

- ・ データ収集方法

問題②では新たなデータ収集は不要のためスキップ。

4-4-3. データ可視化、得たい結論の設定

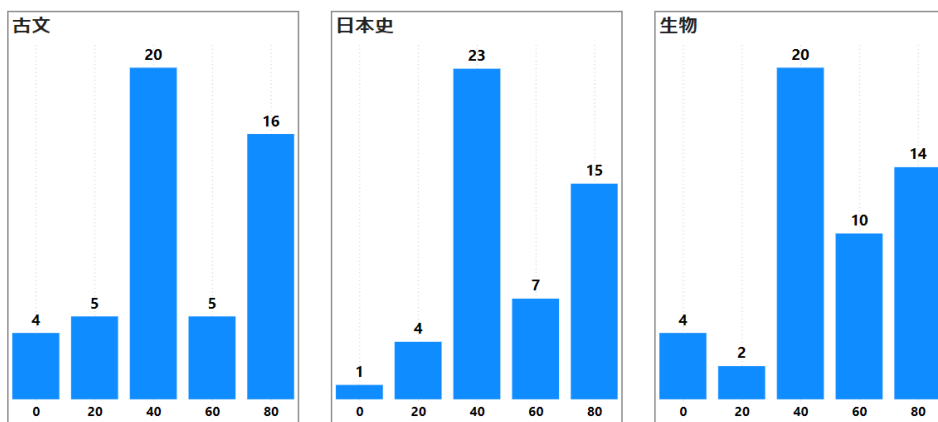
次に「可視化内容」と「可視化した結果から得たい結論」を検討する。

既に「4-3-2. 現状整理」の部分で、「1学期と2学期のテストの点数結果」をもとに現状を整理している。現状整理でまとめた情報から、以下の(1)と(2)について、読み取ることができる。

(1)「4-3-2. 現状整理」で作成したヒストグラムから、各科目「**成績上位**」「**平均点並み**」「**成績下位**」の3グループに分けられることが分かる。

(参考)「4-3-2. 現状整理」から抜粋

・各科目の点数のばらつき【図2】

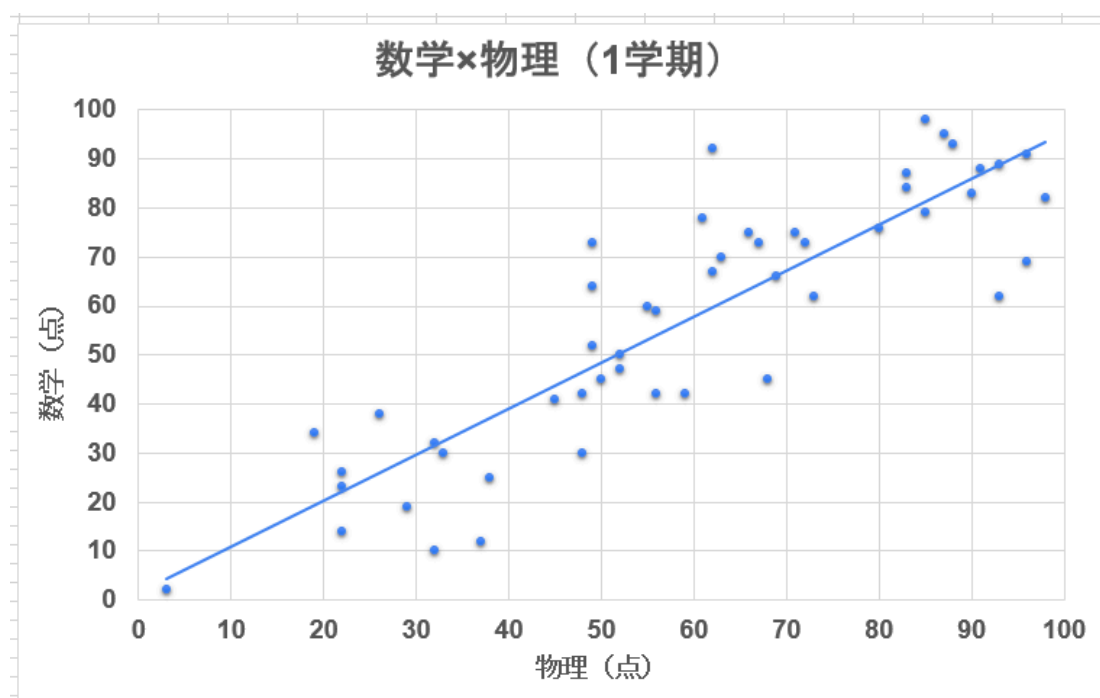


(2)「4-3-2. 現状整理」で作成した散布図から、「**文系科目**」と「**理系科目**」で点数に相関があることが分かる。

(参考)「4-3-2. 現状整理」から抜粋

・テスト結果の相関関係(散布図)【図4】

＜相関性のある散布図(文系と文系、理系と理系)＞



■可視化内容

(1)と(2)より、クラスAの生徒は以下の5つのグループに分けられると仮定し、「1学期と2学期のテストの点数結果」から、クラスAの生徒を以下の5グループに分類できることを可視化する。

【クラスAのグループ分け】

グループ1: 文系理系ともに成績が優秀な人

グループ2: 文系理系ともに成績が平均並みの人

グループ3: 理系の成績だけが低い人

グループ4: 文系の成績だけが低い人

グループ5: 文系理系両方の成績が低い人

ここで定義している「文系」および「理系」の成績については、(2)より文系科目と理系科目で、点数に相関があることが分かった。その

ため、「文系」と「理系」それぞれの科目の合計点をもとにグループ分けを実施することとする。

例)

計画

- ・ 検証方法

手元を与えられている「1学期と2学期のテストの点数結果」を使用して、特性（文系、理系）および成績のレベルに準じたグループ分けを実施する

- ・ 検証データ

「1学期と2学期のテストの点数結果」を使用して検証を実施

- ・ データ収集方法

問題②では新たなデータ収集は不要のためスキップ。

- ・ 可視化内容

「1学期と2学期のテストの点数結果」から、クラスAの生徒を以下の5グループに分類できることを可視化

グループ1: 文系理系ともに成績が優秀な人

グループ2: 文系理系ともに成績が平均並みの人

グループ3: 理系の成績だけが低い人

グループ4: 文系の成績だけが低い人

グループ5: 文系理系両方の成績が低い人

本問題でクラスAの生徒を5グループに分類するにあたり、「標準偏差の68%ルール」を使用することにする。この手法については、データの散らばりを分類する手法として総務省統計局のホームページ等で紹介されており、本仮説検証の中で使用することにした。

「標準偏差の68%ルール」についてもっと詳しく知りたい方は、以下のキーワードで検索してみよう。

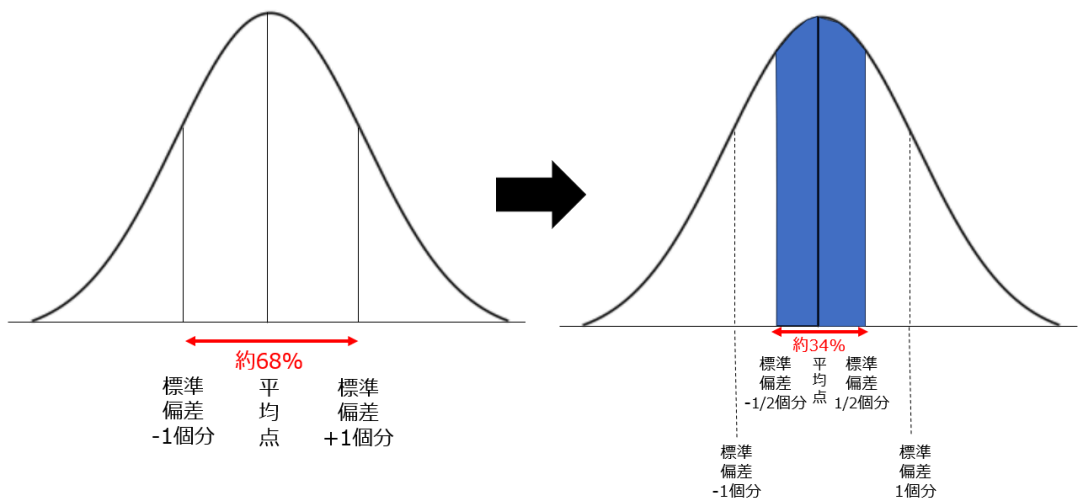
【キーワード】標準偏差の68%ルール

これにより、「1学期と2学期のテストの点数結果」から、クラスAの生徒を5グループに分類できることを可視化できるようになる。

標準偏差の68%ルールとは、「平均値から \pm 標準偏差1個分に含まれるデータは全体の約68%を占める」というルールに基づき、「平均値から \pm 標準偏差1/2個分に含まれるデータは全体の34%を占める」という理論である。

標準偏差の68%ルールにもとづいてグループ分けする手順については、「4-6. データ可視化」で詳しく説明する。

【参考】標準偏差の68%ルールのイメージ図



■可視化した結果から得たい結論

可視化内容をふまえて、「1学期と2学期のテストの点数結果」から、クラスAの生徒を5グループに分類できることを結論として提示していく。

例)

計画

・検証方法

手元に与えられている「1学期と2学期のテストの点数結果」を使用して、特性（文系、理系）および成績のレベルに準じたグループ分けを実施する

・検証データ

「1学期と2学期のテストの点数結果」を使用して検証を実施

・データ収集方法

問題②では新たなデータ収集は不要のためスキップ。

・可視化内容

「1学期と2学期のテストの点数結果」から、クラスAの生徒を以下の5グループに分類できることを可視化

グループ1: 文系理系ともに成績が優秀な人

グループ2: 文系理系ともに成績が平均並みの人

グループ3: 理系の成績だけが低い人

グループ4: 文系の成績だけが低い人

グループ5: 文系理系両方の成績が低い人

・可視化した結果得たい結論

「1学期と2学期のテストの点数結果」から、クラスAの生徒を5グループに分類できることを結論として提示する

【1】～【3】で整理した検証計画の内容をフレームワークの「検証方法」「検証データ」「データ収集方法」「可視化内容」「可視化した結果から得たい結論」にまとめて記載する。

例)

計画

- ・ 検証方法

手元に与えられている「1学期と2学期のテストの点数結果」を使用して、特性（文系、理系）および成績のレベルに準じたグループ分けを実施する

- ・ 検証データ

「1学期と2学期のテストの点数結果」を使用して検証を実施

- ・ データ収集方法

問題②では新たなデータ収集は不要のためスキップ。

- ・ 可視化内容

「1学期と2学期のテストの点数結果」から、クラスAの生徒を以下の5グループに分類できることを可視化

グループ1: 文系理系ともに成績が優秀な人

グループ2: 文系理系ともに成績が平均並みの人

グループ3: 理系の成績だけが低い人

グループ4: 文系の成績だけが低い人

グループ5: 文系理系両方の成績が低い人

- ・ 可視化した結果得たい結論

「1学期と2学期のテストの点数結果」から、クラスAの生徒を5グループに分類できることを結論として提示する

4-5. データ収集

4-5-1. データの取得

4-5-2. データの加工

本問題については、「4-4. 仮説検証の計画」で述べたように、3学期のテストの点数を把握することはできないため、手元にある「1学期と2学期のテストの点数結果」を使用して検証を実施する。そのため、データ収集の手順は省略する。

4-6. データ可視化

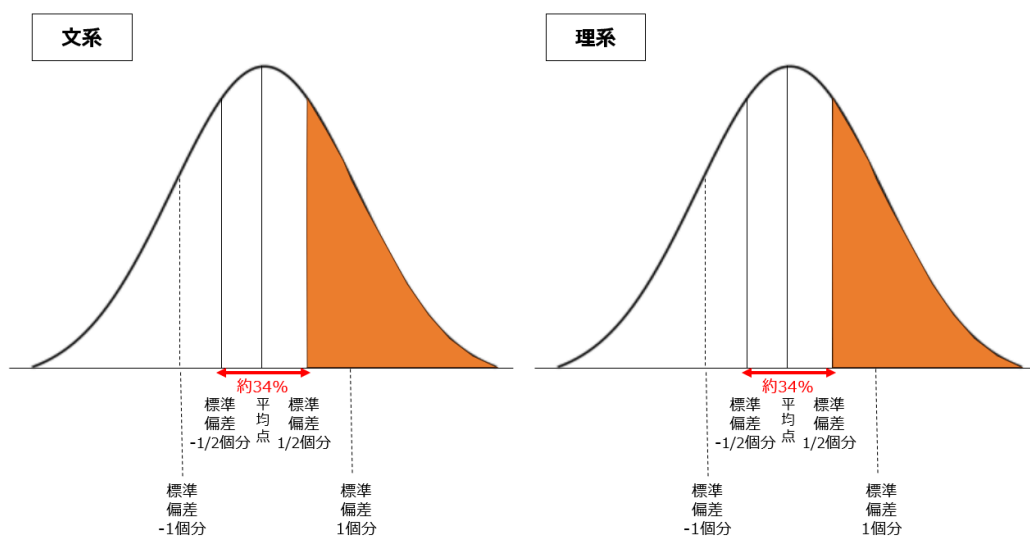
4-6-1. データの可視化

「4-4. 仮説検証の計画」で検討したように、「1学期と2学期のテストの点数結果」から、クラスAの生徒を5グループに分類できることを、「標準偏差の68%ルール」にもとづいて可視化していく。今回はExcel関数を使って、50名を5つのグループに振り分けていく。

■標準偏差の68%ルールで可視化

【クラスAのグループ分け】について、以下のグループ1～5に分類する。

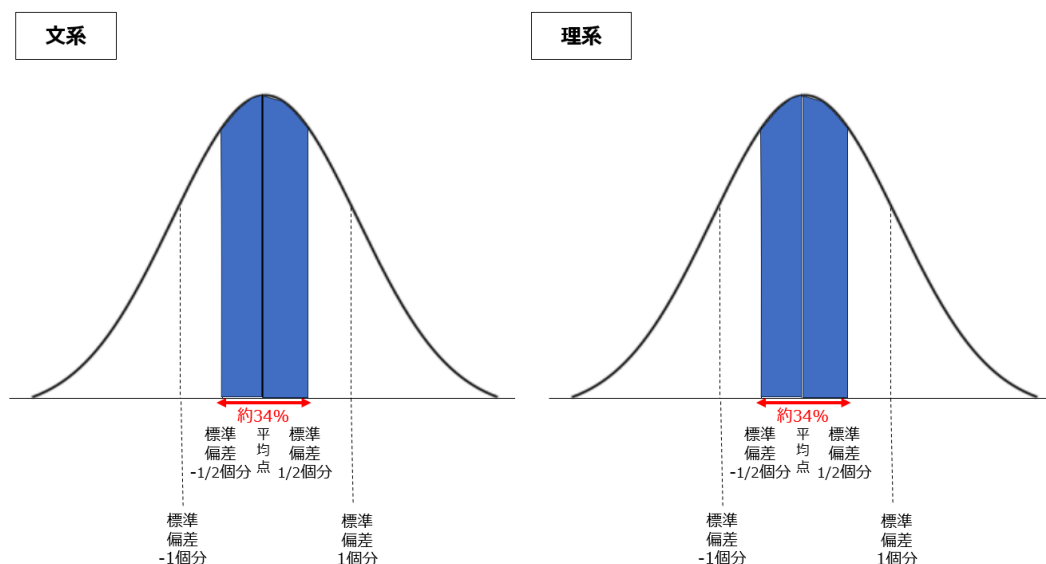
・グループ1: 文系理系ともに成績が優秀な人 (オレンジ色の部分)



文系科目の平均点 \geq 全体の文系科目の平均点 + (標準偏差 \div 2)

理系科目の平均点 \geq 全体の理系科目の平均点 + (標準偏差 \div 2)

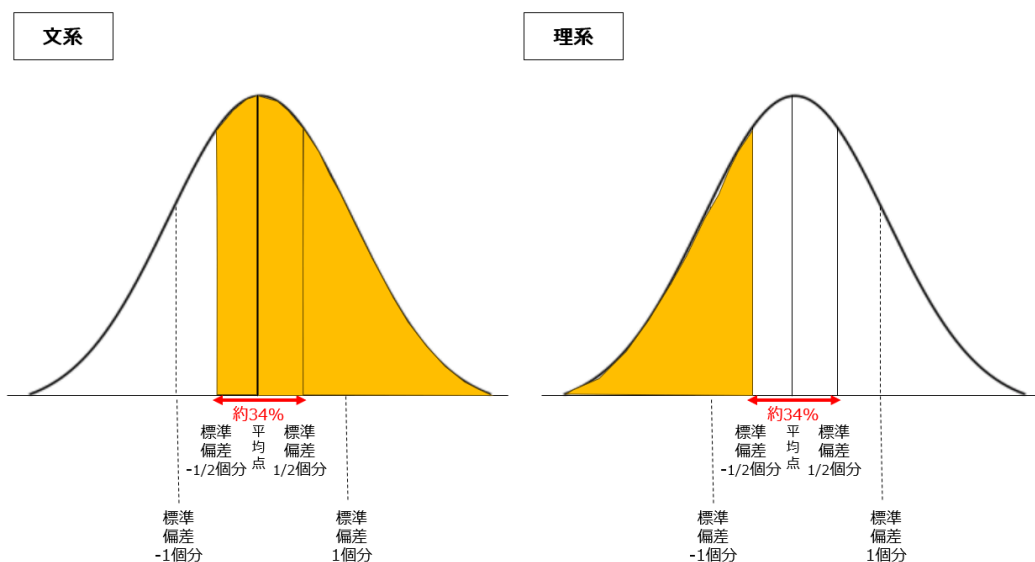
・グループ2: 文系理系ともに成績が平均並みの人 (青色の部分)



全体の文系科目の平均点 - (標準偏差 ÷ 2) ≤ 文系科目の平均点 < 全体の文系科目の平均点 + (標準偏差 ÷ 2)

全体の理系科目の平均点 - (標準偏差 ÷ 2) ≤ 理系科目の平均点 < 全体の理系科目の平均点 + (標準偏差 ÷ 2)

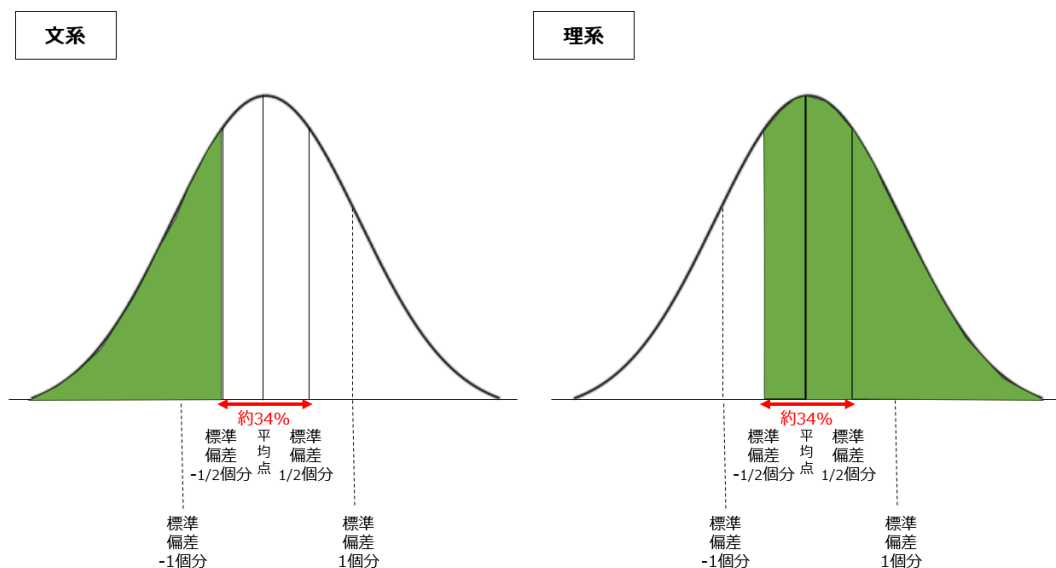
- ・グループ3: 文系の点数が平均並み以上で、理系の点数が低い人
(黄色の部分)



文系科目の平均点 \geq 全体の文系科目の平均点 - (標準偏差 \div 2)

理系科目の平均点 $<$ 全体の理系科目の平均点 - (標準偏差 \div 2)

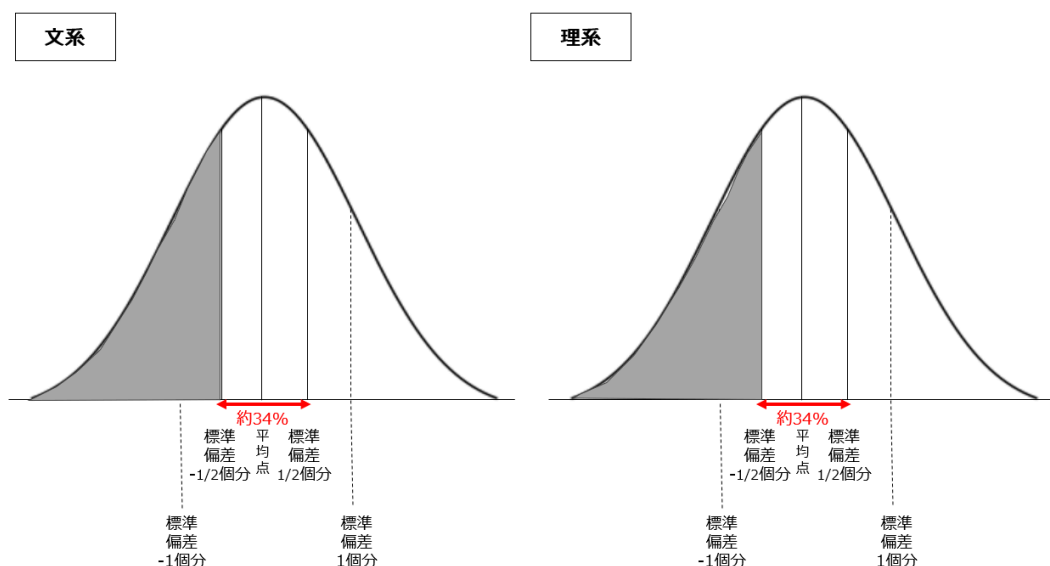
- ・グループ4: 文系の点数が低く、理系の点数が平均並み以上の人
(緑色の部分)



文系科目の平均点 < 全体の文系科目の平均点 - (標準偏差 ÷ 2)

理系科目の平均点 ≥ 全体の理系科目の平均点 - (標準偏差 ÷ 2)

・グループ5: 文系理系両方の点数が低い人(灰色の部分)



文系科目の平均点 < 全体の文系科目の平均点 - (標準偏差 ÷ 2)

理系科目の平均点 < 全体の理系科目の平均点 - (標準偏差 ÷ 2)

上記のグループ分けのルールに従って、Excel関数を使って、クラスAの50名を5つのグループに振り分ける。

Excelで平均点や標準偏差を求める関数式については、【キーワード】で検索し調べていただきたい。

【キーワード】Excel関数 平均、Excel関数 標準偏差

■データの可視化

標準偏差の68%ルールに基づいて、クラスAの生徒を5グループに可視化した結果は以下のとおりである。(一部抜粋)

Excel関数によって、「グループ分け」列(赤枠参照)を作成し、1学期と2学期のテストの点数を、生徒ごとにグループ分けを実施している。

例) 1学期の生徒ごとのグループ分け

名前	現代文	古文	数学	化学	生物	物理	日本史	世界史	政治経	英語	理系平均	文系平均	グループ分け
天照 愛之助	22	58	73	56	59	62	41	23	53	34	62.5	48.1	グループ4
綾小路 あい	98	61	48	60	59	30	63	45	68	82	49.3	61.4	グループ2
雷 功	58	47	49	61	60	73	54	78	73	53	60.8	60.6	グループ2
印宮 伊織	86	95	56	49	56	42	75	63	67	86	50.8	67.5	グループ2
有為橋 海太郎	67	49	68	93	61	45	52	42	70	41	66.8	58.8	グループ2
円城寺 絵見留	53	42	45	59	56	41	68	55	77	55	50.3	55.1	グループ2
応地聖 織戸	44	71	49	62	71	64	59	57	71	78	61.5	62.6	グループ2
女坂 織華	96	91	62	90	81	92	75	93	88	98	81.3	86.6	グループ1
海勢頭 海星	22	26	3	25	11	2	24	51	43	4	10.3	21.1	グループ5
甲楽城 可奈子	57	47	90	80	92	83	46	42	51	85	86.3	67.3	グループ5

【1】で実施したデータの可視化をもとに、1学期と2学期で、グループごとの人数集計をまとめる。

例) 1学期: グループ1～5の人数分布

1学期	グループ	説明	人数
	1	文系理系ともに成績が優秀な人	12
	2	文系理系ともに成績が平均並みの人	16
	3	文系の点数が平均並み以上で、理系の点数が低い人	6
	4	文系の点数が低く、理系の点数が平均並み以上の人	4
	5	文系理系両方の点数が低い人	12

上記の1～5グループに所属する生徒1人1人に対して、所属するグループの分類分けであっているのか確認していくため、各科目のテストの点数結果を確認する作業を実施。

ここでは、元々配布されている、1学期、2学期の10科目のテストの点数結果(実践編_クラス成績向上.csv)をもとに確認する。

4-6-2. データの分析(参考情報)

本問題で求められる「データ分析力」のスキルレベルは4である。「02_データ利活用トレーニングブック_シン・KKDスキル一覧」に記載しているが、レベル4の定義は「レベル3(データからグラフを作って可視化を行える)に加え、グラフから特徴を捉えて結論・考察を整理できる」状態に到達することを目標としている。

データの可視化が実施できるようになり、更にスキルアップを目指したい人は、以降に記載する高度なデータ分析手法にチャレンジしてほしい。

【応用編】

今回の手法だと、文系および理系の合計点でグループ分けを実施しているため、文系の中での科目の優劣を見落としてしまう可能性がある(例えば、現代文が高得点、古文が低得点の場合、合計して平均すると、平均点並みの成績になるため、古文の成績が低い事実を見落としてしまう)。

そのため、より高度な分析手法「主成分分析」や「クラスター分析」を活用すると、上記の課題を解消した上でのグループ分けが可能になる。

「主成分分析」と「クラスター分析」の詳細な分析方法については、**【キーワード】**で検索し調べていただきたい。

【キーワード】主成分分析、クラスター分析

【ヒント】

「主成分分析」と「クラスター分析」を実施する場合の流れは以下のとおりである。

手順①: 主成分分析により、テストの点数に関係している成分をもとめる。

手順② 主成分分析により分かった成分をもとに、互いに似たものを集めて集落(クラスター)を作り、対象を分類していくクラスター分析を実施する。

⇒手順①②より、Excel関数で実施したグループ分けと同様に、クラスAの生徒が5つのグループに分けられることを確認する。

4-7. 結論・考察

4-7-1. 可視化結果と仮説を確認

①可視化結果を記載

「4-6. データ可視化」から読み取った結論をフレームワークに記載する。

②仮説との突合

立てた仮説と可視化した結果を照らし合わせる。

仮説が合っていた場合：

結論・考察を記載し、シン・KKDを完了する。

仮説が間違っていた場合：

結果を踏まえて「①仮説立て」から再度実施する。

○フレームワークに書いてみよう！

		仮説検証の計画	結論・考察
問い	アクション	計画	可視化結果
	→	<ul style="list-style-type: none">・検証方法・検証データ・データ収集方法・可視化内容・可視化した結果得たい結論	→
	→	<ul style="list-style-type: none">・検証方法・検証データ・データ収集方法・可視化内容・可視化した結果得たい結論	→
	→	<ul style="list-style-type: none">・検証方法・検証データ・データ収集方法・可視化内容・可視化した結果得たい結論	→

「因果関係を確認できた(→仮説が正しかった)」場合は、終了となる。「因果関係を確認できなかった(→仮説が正しくなかった)」場合は、ネクストアクションを実施する。ネクストアクションを実施するには、「仮説立て」から「結論・考察」までのプロセスを繰り返し実践することが必要である。

■可視化結果

クラスAの生徒ごとの各科目の点数結果を確認したところ、分類した1～5グループに分けられていることが確認できた。

学力および特性に応じた補講を行うことで、クラスの平均点を向上できると考察できる。具体的な内容を以下に記載する。

【クラスAのグループごとの補講内容】

グループ1: 文系理系ともに成績が優秀な人

→応用問題を解く力を身に着け、高得点を維持させる

グループ2: 文系理系ともに成績が平均並みの人

→基礎もやりながら、応用問題にも挑戦していく

グループ3: 理系の成績だけが低い人

→理系の基礎を重点的に身に着ける

グループ4: 文系の成績だけが低い人

→文系の基礎を重点的に身に着ける

グループ5: 文系理系両方の成績が低い人

→理系および文系の基礎を重点的に身に着ける

フレームワークに分析結果を記載する。

■仮説との突合

本問題では、3学期のテストの点数を確認することはできないので、仮説との突合は省略する。

例)

可視化結果

生徒を5グループに分類することができた。以下のように、学力および特性に応じた補講を行うことで、クラスの平均点を上げることができる考える。

- ・グループ1: 文系理系ともに成績が優秀な人→応用問題を解く力を身に着け、高得点を維持させる
- ・グループ2: 文系理系ともに成績が平均並みの人→基礎もやりながら、応用問題にも挑戦していく
- ・グループ3: 理系の成績だけが低い人→理系の基礎を重点的に身に着ける。
- ・グループ4: 文系の成績だけが低い人→文系の基礎を重点的に身に着ける。
- ・グループ5: 文系理系両方の成績が低い人→理系および文系の基礎を重点的に身に着ける。

以上で「クラス成績向上」は終了とする。



お疲れ様でした！！

5. 健康寿命延伸の検討

(難易度: ★★★)

5-1. 内容

.....

あなたは健康寿命延伸プロジェクトの責任者に任命された。

厚生労働省をはじめとする国が収集したデータを用いて、健康寿命を延伸する(*)ためにはどうしたら良いかを検討する。

(*)本問題では平均寿命の延伸は考慮せず、あくまで健康寿命を延伸するためにはどうしたら良いかを考えよう。



5-2. 目標スキルレベル

この演習問題は、以下の状態になることを目標としている。

- ・論理的思考を用いて問題から課題を検討できるようになる
- ・仮説検証に必要なデータを自分で考え、オープンデータを利用してデータ収集できるようになる

演習問題を解くことで、以下のスキルレベルに到達することを目指してほしい。

スキル	到達レベル
問題解決プロセス / 仮説設定プロセスの理解	★★★★
論理的思考能力	★★★★
データ構造の理解	★★★★
データ収集力	★★★★
データ加工技術	★★★★
データ分析力	★★★★

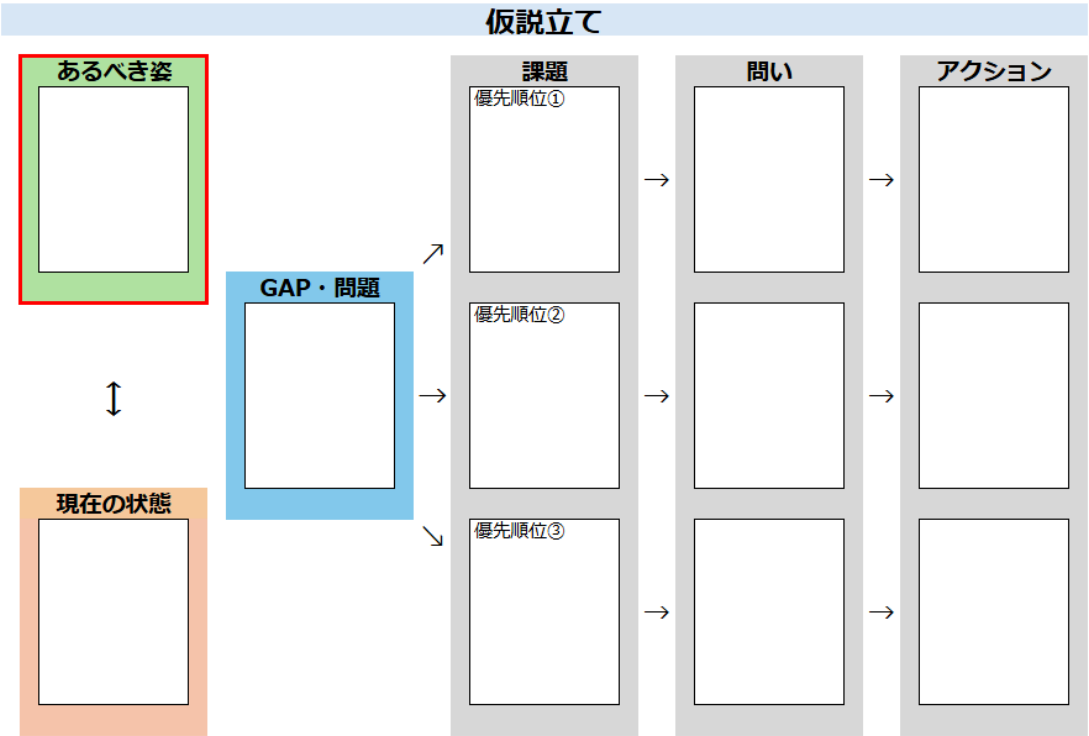
※詳細は「02_データ利活用トレーニングブック_シン・KKDスキル一覧.xlsx」を参照

5-3. 仮説立て

5-3-1. 理想(あるべき姿)の設定

はじめに、成し遂げたい目標を「あるべき姿」として設定する。

○フレームワークに書いてみよう！



例)

今回は、健康寿命を平均寿命と同等になるように延伸すること
あるべき姿として設定する。

あるべき姿

健康寿命を平均寿命と同等になるよ
うに延伸する
男：81.41歳
女：87.45歳

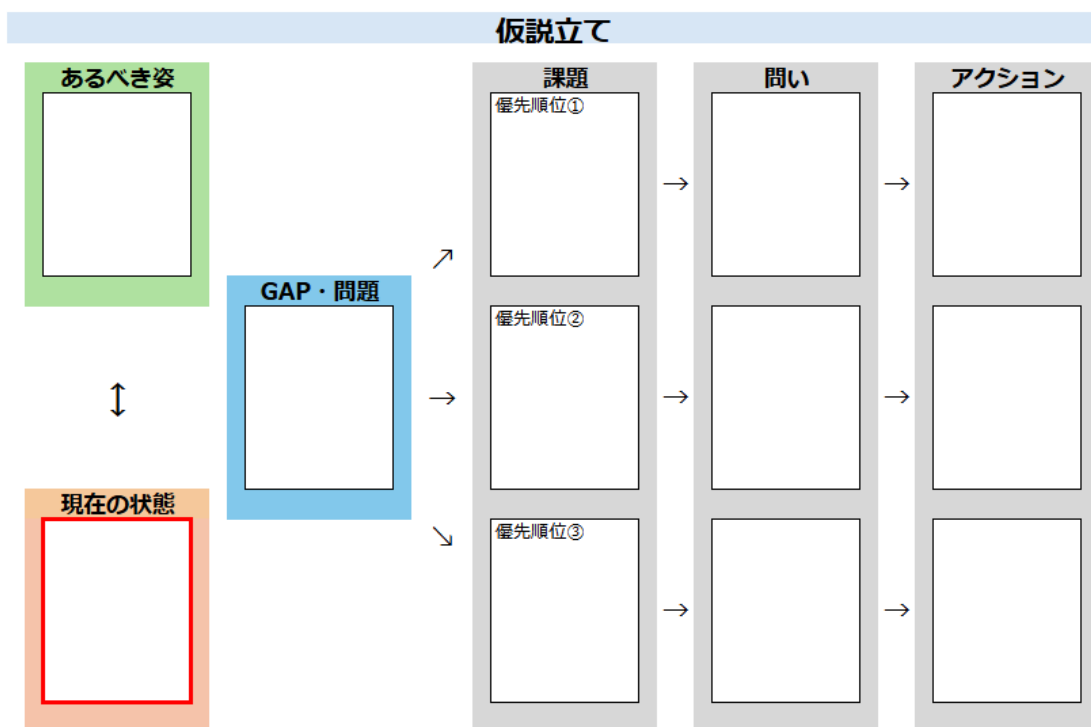
- ・いつ : 10年後まで
- ・どこで : 日本
- ・だれが : 国
- ・何を : 健康寿命
- ・なぜ : ー
- ・どのように : ー

5-3-2. 現状整理

①現在の状態(現状)を記載

日本の健康寿命について、現状分かっていることをe-Statや既存の研究・文献から整理する。

○フレームワークに書いてみよう！



②現状に対するデータ分析を実施

健康寿命延伸の要因は多数存在しており、基礎知識が乏しい状態では、仮説立てやロジックツリー作成すら難しい。

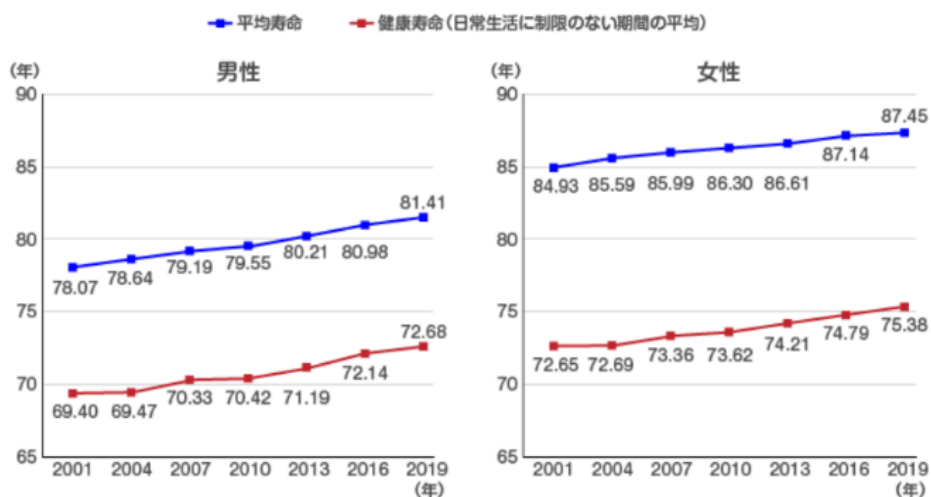
今後の検証を効率良く進める上で必要な基礎知識を吸収するためにも、現状分析を丁寧に行う。健康寿命延伸には既存の研究・文献が数多く存在するため、基礎知識の網羅に活用する。

また、研究・文献から得ることができない情報については、e-Statを活用しデータの可視化を行う。

例)

既存の研究・文献から現状整理を行う場合

図1.平均寿命と健康寿命の推移^[2]



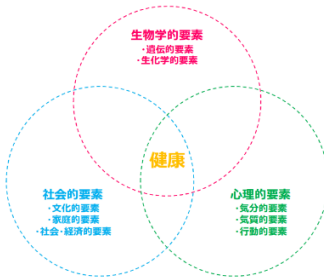
平均寿命と健康寿命 | e-ヘルスネット(厚生労働省)

*サリバン法

毎年必ず10万人が誕生する状況を仮定し、そこに年齢別の死亡率と、年齢別の「健康・不健康」の割合を与えることで、「健康状態にある生存期間の合計値(健康な人の定常人口)」を求め、これを10万で除して健康寿命を求めます。なお、我が国の現行指標では、簡易生命表から5歳階級別の定常人口、国民生活基礎調査から5歳階級別の「健康・不健康」の割合を得て、「健康な人の定常人口」を求めています。

健康寿命の定義と算出方法 | e-ヘルスネット(厚生労働省)

図表 5 「健康」の構造（生物心理社会モデル）



(出所) Engel (1977) "The Need for a New Medical Model: A Challenge for Biomedicine"
より作成

図表 21 健康寿命の延伸に向けたフロー



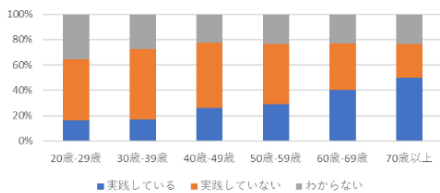
(出所)「健康日本 21（第二次）」の概念図を基に作成

健康増進法施行規則等の一部を改正する省令

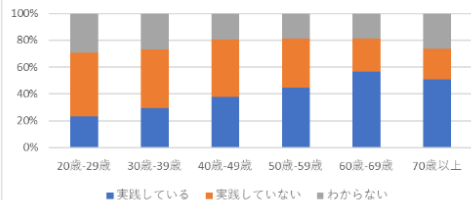
例)

e-Statのデータを可視化し、現状整理する場合

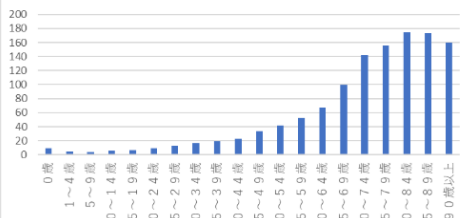
健康寿命を延伸するために良い生活習慣を実践している者の割合（2011・男）



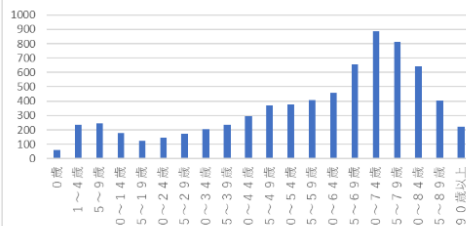
健康寿命を延伸するために良い生活習慣を実践している者の割合（2011・女性）



入院患者数(2020)



外来患者数(2020)



例)

現在の状態

平均寿命と健康寿命に男女共に差がある。

(2019)

【性別：平均寿命：健康寿命】

男：81.41歳：72.68歳(8.73歳差)

女：87.45歳：75.38歳(12.06歳差)

健康寿命の要因は大きく3種類に分かれる。

生物学的要素

社会的要素

心理的要素

健康寿命を延伸させるためには、

- ①死亡率の低下
- ②不健康割合の低下

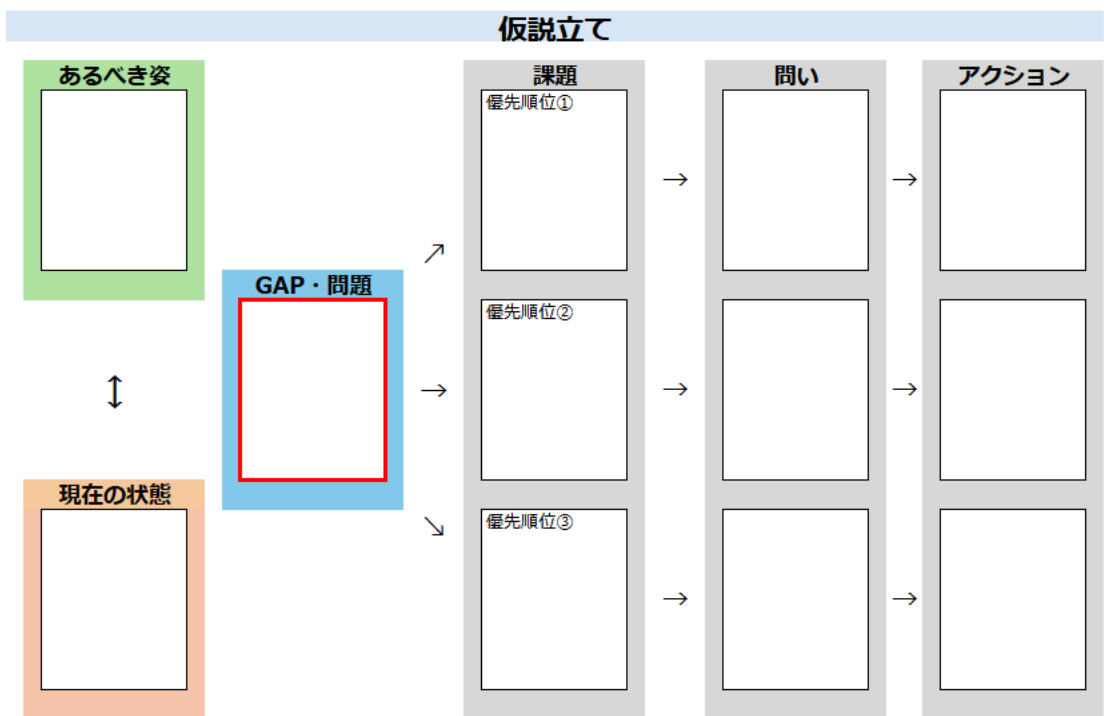
の2つがある。

70歳までは、年を取るほど健康寿命を延伸させるために活動を行っている人の割合が増える。
70歳以降は減っている。

5-3-3. GAP・問題の設定

設定した「あるべき姿」と整理した「現状」の差分（「コト」や「状況」）を「GAP・問題」として書き出す。

○フレームワークに書いてみよう！



ここでは、シンプルに「あるべき姿」と「現状」の差を書き出す。曖昧な表現、憶測や主観的な表現は避ける。

例)

あるべき姿

健康寿命を
平均寿命と同等になるように延伸する

問題

平均寿命と健康寿命には男女で以下の差がある。
男性：8.73歳
女性：女：12.06歳

現状

平均寿命と健康寿命は以下の通り。
男性 平均寿命：81.41歳 健康寿命：72.68歳
女性 平均寿命：87.45歳 健康寿命：75.38歳

GAP・問題

現在の平均寿命を維持した場合、健康寿命を以下くらい伸ばす必要がある

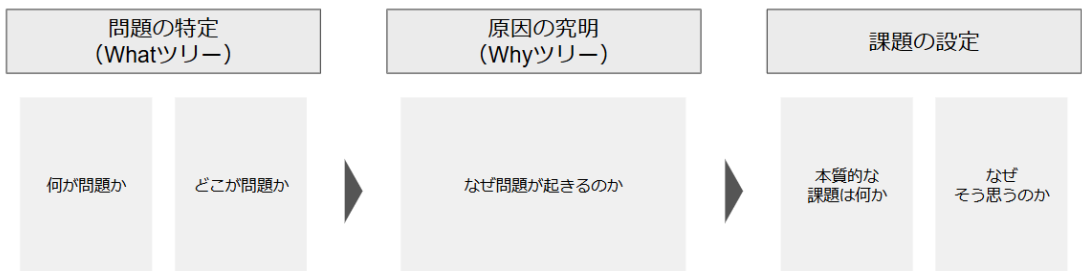
男：8.73歳
女：12.06歳

5-3-4. 課題の設定

導き出された「問題」が、どこで起きているのかを明らかにし、「課題」に落とし込む。

その際、ロジックツリーを用いて、物事の要素を分解して網羅的に洗い出すことで問題の発生個所を特定していく。MECEを意識して、問題の全体像や関係性を細かく分解していく。

○ロジックツリー記載用シートに書いてみよう！

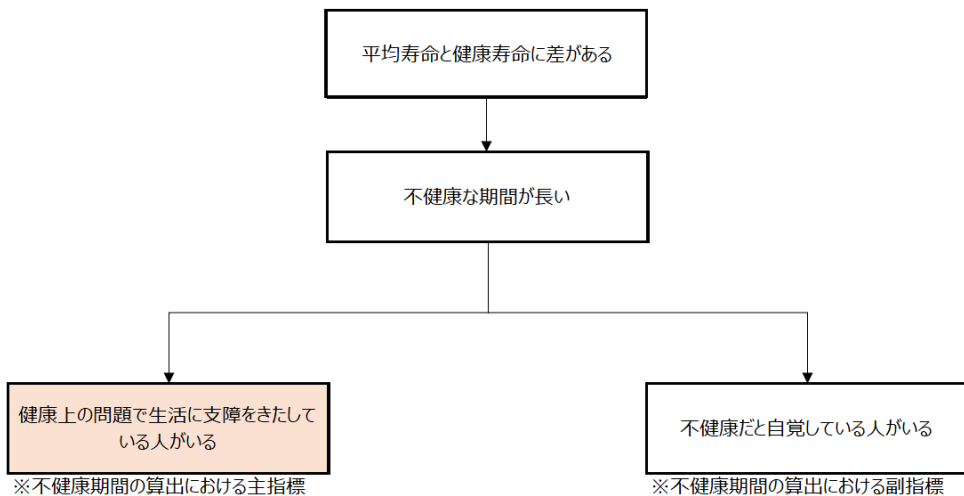


記載する際は、「問題の特定」「原因の究明」「課題の設定」を意識するとよい。

①ロジックツリーの作成

ロジックツリーを作成するときは、設定したGAP・問題から原因を究明する。

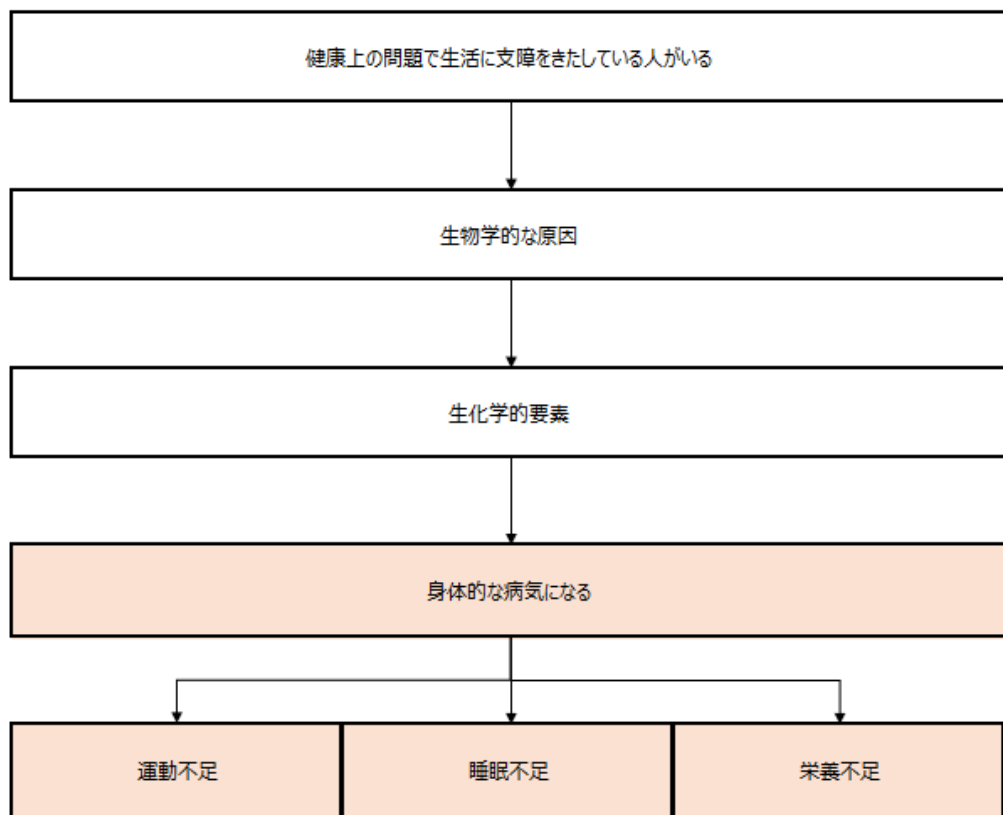
例) Whatツリー



例) Whyツリー

※オレンジの箇所を課題として設定

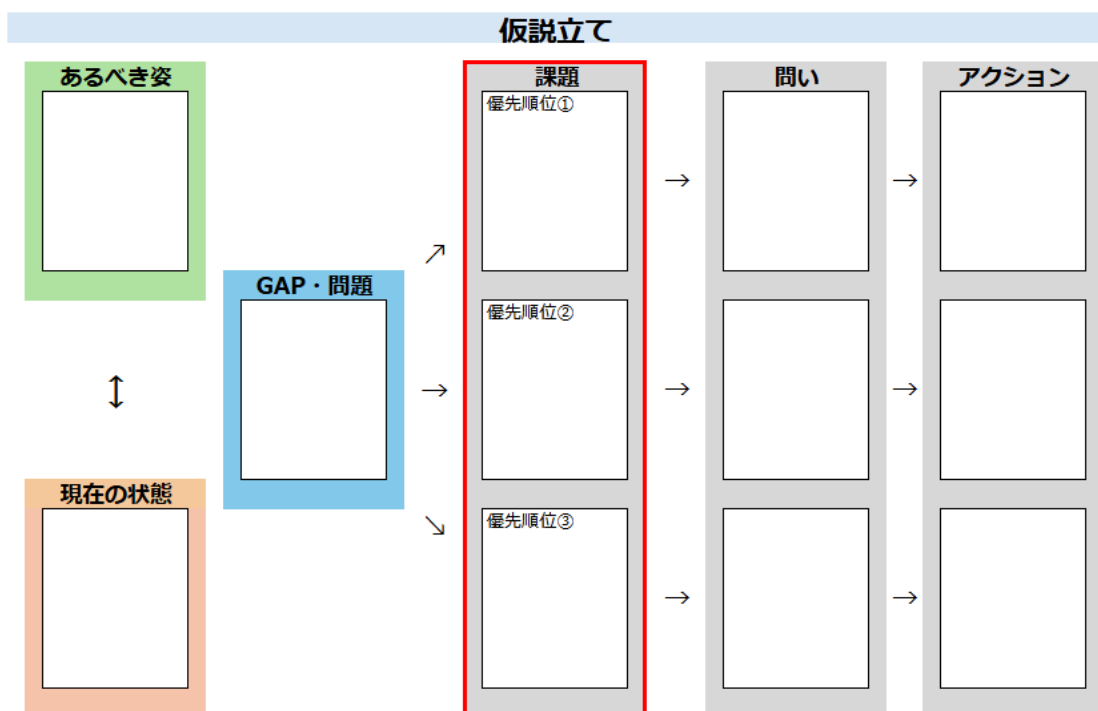
※ツリーの一部を抜粋



②課題を記載

原因究明のステップで作成したロジックツリーの内容から、本質的な課題は何なのか抽出していく。

○フレームワークに書いてみよう！



例)

課題

優先順位①

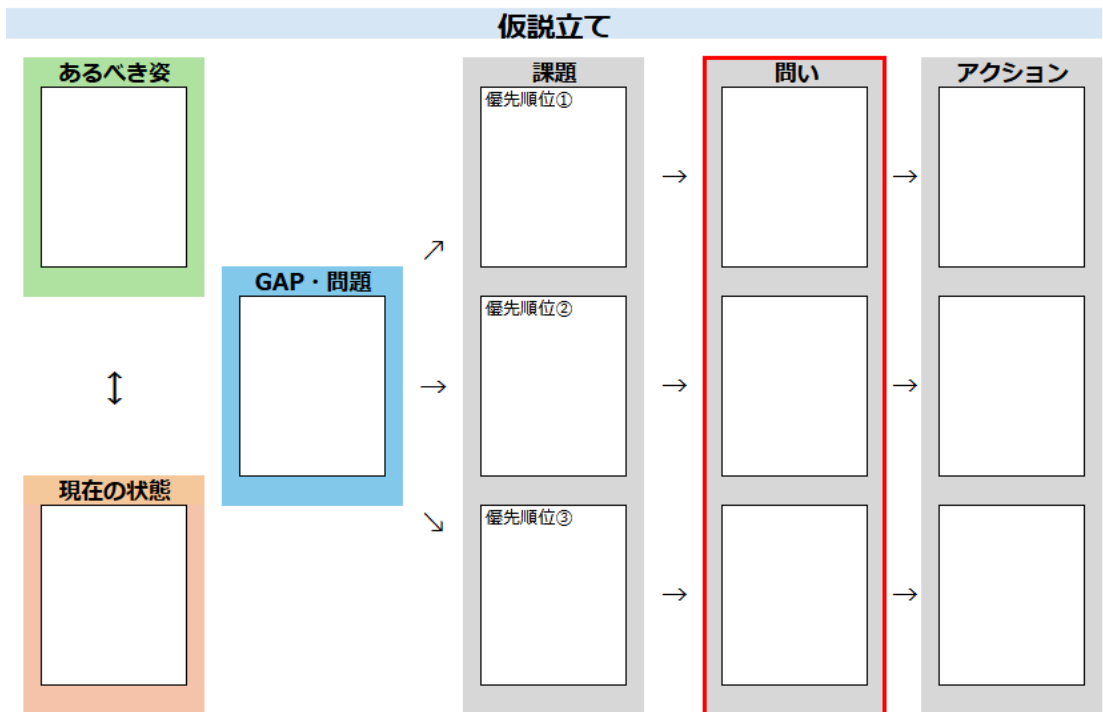
身体的な病気を患っていることにより、健康を維持できない

なお、データ可視化の結果から読み取れる課題であるか(それが課題であるエビデンスがあるか)を確認する。

5-3-5. 問いの設定

「あるべき姿」と現状とのギャップを考えながら「問い」を考える。

○フレームワークに書いてみよう！



例)

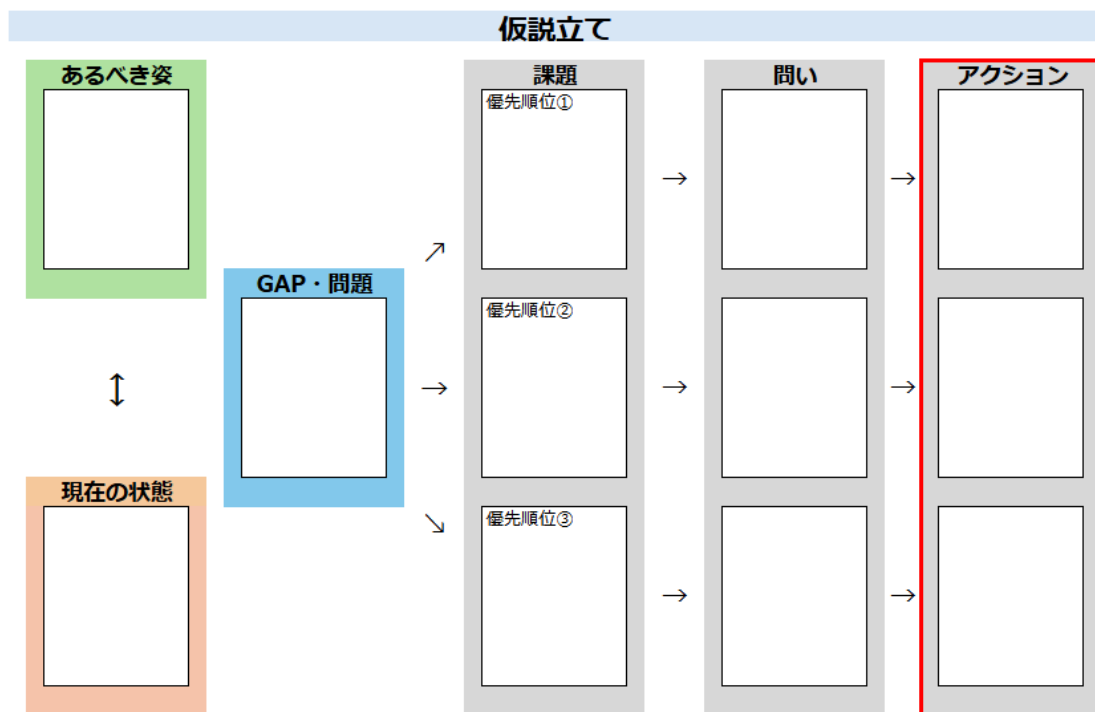
問い

日本人の健康に特に悪影響のある身体的な病気を減らすことで、健康寿命を延ばすことができるのではないか。

5-3-6. アクションの設定

「5-3-5. 問いの設定」で記載した、課題を解決するための仮説として「アクション」を記載する。「アクション」の記載にあたり、「誰が、いつまでに」が明確になるように記載を行う。

○フレームワークに書いてみよう！



例)

アクション

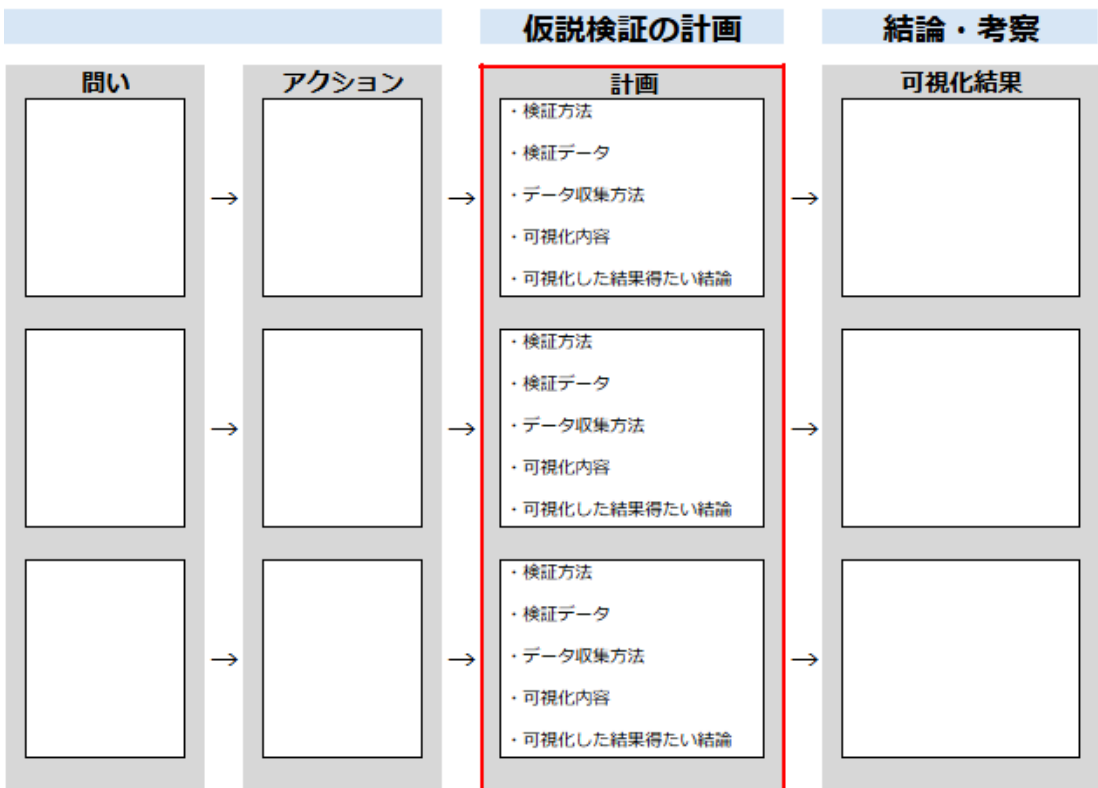
身体的な病気と健康寿命の関連性を調査し、健康寿命に悪影響を及ぼす病気を特定する。

5-4. 仮説検証の計画

5-4-1. 検証方法の設定

「アクション」により課題が解決するのか、を検証する方法を検討し、その内容を記載する。

○フレームワークに書いてみよう！



■検証方法

本問題では、「アクション」にデータ可視化を記載することが多いため、「アクション」を満たす検証方法を記載する。

例)

計画	
・検証方法	傷病数の推移と健康年齢の推移を比べる
・検証データ	2001年～2019年の健康年齢 2001年～2019年の傷病ごとの人数
・データ収集方法	健康年齢は厚生労働省が公開 e-Statより国民生活基礎調査の健康の分類から収集
・可視化内容	散布図 バブルプロット
・可視化した結果得たい結論	特定の傷病と健康年齢に相関がある

5-4-2. データ収集方法の設定

① 検証データを記載

検証に必要な「データ」を検討し、記載する。

② データ収集方法を記載

検証に必要なデータの収集方法を検討し、記載する。

可視化に必要なデータのサンプル数が足りない場合がある。

そのため、データを男女別や都道府県別などに分割し、サンプル数を確保することも視野に入れる。

例)

計画	
・検証方法	傷病数の推移と健康年齢の推移を比べる
・検証データ	2001年～2019年の健康年齢 2001年～2019年の傷病ごとの人数
・データ収集方法	健康年齢は厚生労働省が公開 e-Statより国民生活基礎調査の健康の分類から収集
・可視化内容	散布図 バブルプロット
・可視化した結果得たい結論	特定の傷病と健康年齢に相関がある

5-4-3. データ可視化、得たい結論の設定

次に「可視化内容」と「可視化した結果から得たい結論」を検討する。

■可視化内容

設定した「検証方法」に基づいて、データをどのように可視化するのかを設定する。

グラフによってデータの見え方が変わってくるため、後述の「可視化した結果得たい結論」と合わせて慎重に設定する。

■可視化した結果から得たい結論

「可視化内容」を踏まえて得たい結論を記載する。

例)「可視化内容」には、傷病数と健康寿命の相関関係を明らかにするために必要な可視化方法を記載する。「可視化した結果得たい結論」には、5-3-6で設定した「アクション」を満たす内容を記載する。

計画	
・検証方法	傷病数の推移と健康年齢の推移を比べる
・検証データ	2001年～2019年の健康年齢 2001年～2019年の傷病ごとの人数
・データ収集方法	健康年齢は厚生労働省が公開 e-Statより国民生活基礎調査の健康の分類から収集
・可視化内容	散布図 バブルプロット
・可視化した結果得たい結論	特定の傷病と健康年齢に相関がある

5-5. データ収集

5-5-1. データの取得

検証で利用するデータの収集を行う。本問題では、オープンデータ(e-Stat)と、既存の研究論文からデータの可視化結果を収集する。

5-5-2. データの加工

①加工用データの準備

「【1】データの取得」にて収集したデータを、Excelなどのソフトを利用して説明文など不要な行を削除する。

また、計画どおりのデータ収集が難しい場合は、「5-4.仮説検証の計画」の見直しを行う。

例)ヘッダーなど可視化に不要な情報を取り除く。

今回は総数を利用するため、年齢ごとのデータも取り除く。

■ 加工前

統計名：	令和元年国民生活基礎調査 健康 全国編									
表番号：	98									
表題：	[通院者の状況（第98表～第107表）] 通院者数，年齢（5歳階級）・最も気になる傷病・性別									
実施年月：	2019年 -									
表章項目：	1660	通院者数【千人】								
調査年度：	13	2019年(令和元年)								
		/性_002 コ								
		/性_002								
		/傷病（複	110	120	130	140	150	160	170	
		/傷病（複	通院者数	糖尿病	肥満症	高コレステ	甲状腺の病	その他のこ	認知症	-
110	総数		50,045	3,784	103	2,012	712	1,650	488	
120	0～4歳		605	-	-	0	2	0	-	-
130	5～9歳		855	1	1	-	3	13	-	-
150	10～14歳		882	1	1	0	4	29	-	-
170	15～19歳		706	6	1	1	13	48	-	
200	20～24歳		647	7	0	1	14	67	-	-
210	25～29歳		824	14	1	7	16	103	-	-
240	30～34歳		1,156	20	1	9	39	146	0	
250	35～39歳		1,608	48	6	23	40	174	-	

■ 加工後

		110	120	130	140	150	160	170	180
年齢(5歳階	通院者数	糖尿病	肥満症	高コレステ	甲状腺の病	その他のこ	認知症	パーキンソン	
総数	50,045	3,784	103	2,012	712	1,650	488	176	

②データを整形

収集したデータがそのまま利用できるケースは少ない。理論編に記載している「欠損値」や「外れ値」、「表記ゆれ」などの対応を行う。

例)年代ごとにデータ収集の単位が異なるため、最新データを基準に統合または分割を行う。

例えば、2019年の「耳の病気」は、2004年以前では「中耳炎」「難聴」に分割されている。

2019	2016
耳の病気	耳の病気
2004	2001
中耳炎	中耳炎
難聴	難聴

③データにカテゴリーを付与

本問題ではカテゴリー付与を行わなかったため、省略する。

単一のデータでは可視化が難しい場合、データの結合を行い、可視化を試みる。

※可視化が目的であるため、書式の統一は行っていない

図表1-2-6 平均寿命と健康寿命の推移		
(年)		
男性	平均寿命	健康寿命
2001年	78.07	69.40
2004年	78.64	69.47
2007年	79.19	70.33
2010年	79.55	70.42
2013年	80.21	71.19
2016年	80.98	72.14

(年)		
女性	平均寿命	健康寿命
2001年	84.93	72.65
2004年	85.59	72.69
2007年	85.99	73.36
2010年	86.30	73.62
2013年	86.61	74.21
2016年	87.14	74.79

資料: 平均寿命については、2010年につき厚生労働省政策統括官付参事官付人口動態・保健社会統計室「完全生命表」、他の年につき「簡易生命表」、健康寿命については厚生労働省政策統括官付参事官付人口動態・保健社会統計室「簡易生命表」、「人口動態統計」、厚生労働省政策統括官付参事官付世帯統計室「国民生活基礎調査」、総務省統計局「人口推計」より算出。

107

■ 元データ(人口)

※全体の健康寿命算出のため

統計名：	人口推計 各年10月1日現在人口 平成27年国勢調査基準 統計表								
表番号：	001								
表題：	年齢（各歳），男女別人口及び人口性比－総人口，日本人人口								
実施年月：	-								
市区町村別：	-								
全国：	00000	全国							
時間軸（年）	1201	2019年10月1日現在							

年齢各歳	年齢各歳	/人口 /人口 /男女別・ /男女別・	001				002			
			総人口				日本人人口			
			001	002	003	004	001	002	003	004
			女計【千人】	男【千人】	女【千人】	比【女性】	女計【千人】	男【千人】	女【千人】	比【女性】
01000	総数		126,167	61,411	64,756	94.8	123,731	60,208	63,523	94.8
01001	0歳		894	458	436	105.2	877	450	427	105.2
01002	1歳		941	482	459	105.1	925	474	451	105.1
01003	2歳		962	493	469	105.1	946	484	461	105.0
01004	3歳		1,001	514	487	105.6	985	506	479	105.6
01005	4歳		961	490	470	104.3	946	483	463	104.3
01006	5歳		975	499	476	104.9	959	491	468	104.9
01007	6歳		1,012	517	495	104.5	997	510	488	104.5

[人口推計](#) | [データベース](#) | [統計データを探す](#) | [政府統計の総合窓](#)

[□](#)

■ 元データ(傷病)

[illegible]

[国民生活基礎調査 令和4年国民生活基礎調査](#) | [ファイル](#) | [統計](#)

[データを探す | 政府統計の総合窓口](#)

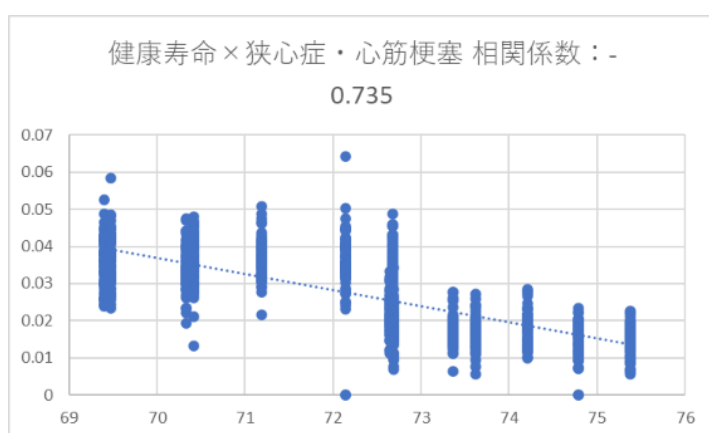
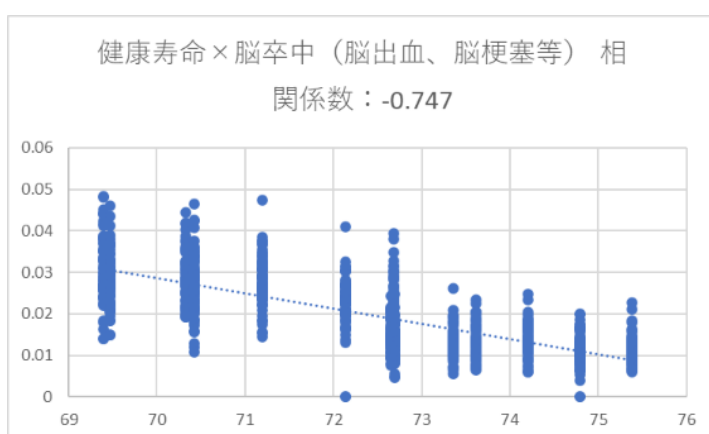
■ 加工データ(健康寿命×人口×傷病)

	全体			糖尿病	肥満症	高コレステロール血症	中脳卒中	その他の心臓病	認知症	パーキンソン病	がん（神経）	眼の病気	耳の病気
(年)	人口（千人）	平均寿命	健康寿命	健康寿命×男	健康寿命×女	健康寿命×男	健康寿命×女	健康寿命×男	健康寿命×女	健康寿命×男	健康寿命×女	健康寿命×男	健康寿命×女
平成13 (2001)	127,291	81.58	71.06	0.0495	0.002	0.0304	0.0097	0.012	0	0	0.0171	0.0309	0.013
16 (2004)	127,687	82.2	71.12	0.0563	0.0025	0.0287	0.0105	0.0167	0	0	0.0178	0.028	0.0125
19 (2007)	127,771	82.67	71.88	0.0553	0.0021	0.0297	0.0112	0.0258	0.0054	0.0027	0.008	0.0346	0.0119
22 (2010)	128,057	83.02	72.06	0.0625	0.0023	0.0408	0.0121	0.0293	0.007	0.0029	0.009	0.0373	0.0112
25 (2013)	127,298	83.5	72.74	0.0709	0.0025	0.0361	0.014	0.0316	0.0094	0.0037	0.0095	0.0391	0.0108
28 (2016)	126,933	84.14	73.5	0.0733	0.0022	0.0375	0.0137	0.031	0.0106	0.0034	0.0101	0.0405	0.0105
令和元 (2019)	126,167	84.51	74.07	0.0756	0.0021	0.0402	0.0142	0.033	0.0098	0.0035	0.0099	0.0422	0.0096
相関係数				0.948	-0.107	0.759	0.937	0.866	0.915	0.835	-0.637	0.943	-0.973

5-6. データ可視化

「5-5. データ収集」にて収集・加工したデータを使って、データを可視化する。必要に応じてデータから算出される指標と合わせて分析を行う。(標準偏差や相関係数など)

例) 健康寿命と最も気になる症状(全体)



5-7. 結論・考察

5-7-1. 可視化結果と仮説を確認

①可視化結果を記載

「5-6. データ可視化」の結果をフレームワークに記載する。各アクション実行後、どのような状態になったかを記載する。

②仮説との突合

立てた仮説と可視化した結果を照らし合わせる。

仮説が合っていた場合：

結論・考察を記載し、シン・KKDを完了する。

仮説が間違っていた場合：

結果を踏まえて「①仮説立て」から再度実施する。

○フレームワークに書いてみよう！



「因果関係を確認できた(→仮説が正しかった)」場合は、終了となる。「因果関係を確認できなかった(→仮説が正しくなかった)」場合は、ネクストアクションを実施する。ネクストアクションを実施するには、「仮説立て」から「結論・考察」までのプロセスを繰り返し実践することが必要である。

例)

可視化結果

結論：

分析結果はグラフシートの健康寿命との負の相関（強い順）の通り。
回答時にたまたまその傷病を患っていた、という可能性があるため、これだけでは健康寿命との関連は不明。

※グラフフォルダ参照

考察：

日常的に不便さを感じやすそうな病気(症状を自覚しやすい病気)ほど、健康寿命と負の相関があるように見える。

最も気になる症状と健康寿命の相関関係を調べることで、健康寿命に直結する疾病を特定することができるのではないか。

以上で「健康寿命延伸の検討」は終了とする。



読了、お疲れ様でした！！



Thanks