

AMAZON QUICKSIGHT ハンズオン

Amazon Web Services Japan

第 2 版 2020/03/11 作成, 最終更新: 2020/04/20

目次

本資料について	4
事前準備と環境	4
ハンズオンの流れ	5
QuickSight の利用を開始する	6
QuickSight にサインアップする	6
事前準備: CSV ファイルをアップロードし、データセットを作成する	9
QuickSight と SPICE	9
CSV ファイルを確認する	9
SPICE に 1 つ目の CSV ファイルをアップロードする	10
データセットを準備 (PREPARE) する	12
ハンズオン 1: 基本的な可視化とクイックフィルタ	18
垂直積み上げ棒グラフの作成	18
ドーナツグラフの作成	22
ピボットテーブルとコンディショナルフォーマットの作成	25
クイックフィルタ機能による絞り込みの実現	30

（オプション演習）折れ線グラフと予測機能（Enterprise Edition が必要）	33
（オプション演習）テーマを切り替える	37
ハンズオン 2：より進んだ可視化	38
ハンズオン 2 を始める前に	38
垂直棒グラフの作成とドリルダウンの実現	40
ビジュアルの調整とコピー	45
コントロールの作成とフィルタ	47
計算フィールド①：シンプルな計算フィールド	52
計算フィールド②：テーブル計算関数（runningSum）	55
（オプション）計算フィールド③：テーブル計算関数（percentDifference）	57
（オプション）計算フィールド④：テーブル計算関数（percentOfTotal）	59
ML インサイトによる異常値検出の設定	61
ML インサイトによるナラティブ	64
ML インサイトによる異常値検出の内容を確認する	66
（オプション）ハンズオン 3：Level-Aware Aggregation	68
Level-Aware Aggregation とは	68
Level-Aware Aggregation (countOver)	69
Level-Aware Aggregation (sumOver)	71
まとめと後片付け	74
補足情報・参考資料（QuickSight）	75

本資料について

事前準備と環境

本資料は Amazon Web Services（以下 AWS）が提供する BI サービス、Amazon QuickSight（以下 QuickSight）を使ってみることで、その機能を体験するハンズオンです。

ハンズオンを実行するには AWS のアカウントが必要になります。まだ AWS アカウントをお持ちでない場合は、以下より登録が可能です。

<https://aws.amazon.com/jp/register-flow/>

ハンズオンは QuickSight のみで構成されていますが、ハンズオンの内容を超えて QuickSight から各種データソース、例えば Amazon S3（以下 S3）、Amazon Athena（以下 Athena）等にアクセスする場合は、最低限それら进行操作できる権限を持つユーザー（IAM ユーザ）を使用してください。

それぞれの サービスは所定の費用がかかります。費用の詳細は以下のホームページでご確認ください。本ハンズオンでは、QuickSight の基本機能は無料枠の範囲（最初の 1 ユーザのみ、SPICE 1GB 未満）で実施できるよう設計されていますが、ML（機械学習）インサイトの異常検知の機能は、別途費用がかかります（ハンズオン 2 の最後の部分）。また、ハンズオン記載以上の処理を実施するとその分費用が発生する可能性がある点にご注意ください。

- QuickSight の料金（最初の 1 ユーザ、SPICE 1GB 分までは無料利用枠あり）

<https://aws.amazon.com/jp/quicksight/pricing/>

- Athena の料金

<https://aws.amazon.com/jp/athena/pricing/>

- S3 の料金

<https://aws.amazon.com/jp/s3/pricing/>

ハンズオンを実行するには、ブラウザ（**Chrome か Firefox の最新版を推奨**します。その他 Safari 7 以降、Edge 最新バージョン、IE11 以降もサポートされます）とインターネットに繋がる環境が必要です。ブラウザと QuickSight サービス間の通信は、HTTPS と WebSockets Secure(wss://)を使用しますので、ファイアーウォール等で防がれていないかご確認ください。

また、利用ブラウザのプラグイン・アドオンに御注意ください。特に自動翻訳のプラグインが QuickSight の動作を妨げる事がありますので事前の停止をおすすめします。

ハンズオンの流れ

最初に QuickSight のアカウント（AWS アカウントとは別です）にサインアップしてデータセットを準備し、その後以下の 2 パートのハンズオンを試すことができます。パート 1 とパート 2 は独立しているため、どちらか一方だけを実施することも可能です。

- ハンズオン：パート 1

パート 1 では、棒グラフやドーナツチャート、ピボットテーブルといった基本的なビジュアルの作成方法に加え、クイックフィルタの設定や、コンディショナルフォーマット（値によって色を変える）を体験します。オプションで折れ線グラフを作成し、予測を追加する方法やテーマ（配色）切り替えを体験します。

- ハンズオン：パート 2

パート 2 ではより進んだ活用方法を体験します。コントロール（GUI）を配置してのフィルタの設定、計算フィールド（関数）の利用、ML インサイトによる異常検知と自動ナラティブの作成を体験します

- ハンズオン：パート 3（オプション）

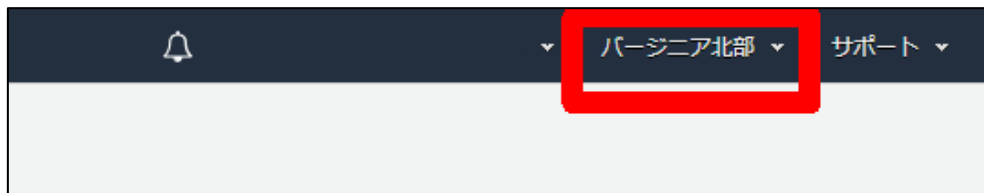
高度な関数の使い方である Level Aware Aggregations (LAA)の利用方法を体験します。

QUICKSIGHT の利用を開始する

本章では QuickSight にサインアップし、利用開始できるまでを説明します。すでに QuickSight を利用されている場合は本章をスキップしてください。

QUICKSIGHT にサインアップする

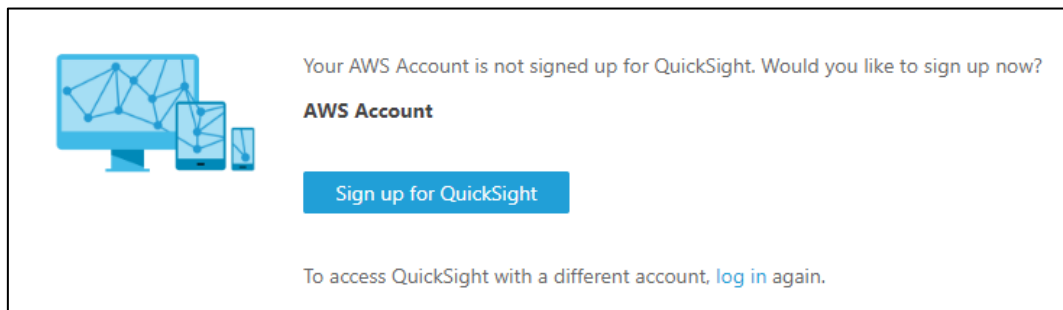
<https://console.aws.amazon.com> にアクセスし、AWS マネジメントコンソールにログインします。AWS マネジメントコンソールでは、右上のメニューから利用リージョンを「バージニア北部」に設定してください。



サービス一覧から **QuickSight** を選択します。検索欄で quick と入力すると絞り込まれます。



初めて利用する場合は以下の画面が表示されますので、“**Sign up for QuickSight**”をクリックします。



クリックすると以下のエディション選択の画面が出ます。

Edition	<input type="radio"/> Standard <input checked="" type="radio"/> Enterprise	
	Standard	Enterprise
First author with 1GB SPICE	FREE	FREE
Team trial for 60 days (4 authors)*	FREE	FREE
Additional author per month (yearly)**	\$9	\$18
Additional author per month (monthly)**	\$12	\$24
Additional readers (Pay-per-Session)	N/A	\$0.30/session (max \$5/reader/month) *...
Additional SPICE per month	\$0.25 per GB	\$0.38 per GB
Single Sign On with SAML or OpenID Connect	✓	✓
Connect to spreadsheets, databases & business apps	✓	✓
Access data in Private VPCs		✓
Row-level security for dashboards		✓
Hourly refresh of SPICE data		✓
Secure data encryption at rest		✓
Connect to your Active Directory		✓
Use Active Directory Groups ***		✓
Send email reports		✓

どちらのエディションでも最初の 1 ユーザ無料で利用できるため、ここでは **Enterprise Edition** を選択します。ハンズオンで ML インサイトの異常検知を試す場合には Enterprise Edition が必要です。なお Standard Edition を選択しても後から Enterprise Edition にアップグレード可能です。選択したら、画面下方にある Continue を押します。

☒ Use Role Based Federation (SSO)
☐ Use Active Directory

QuickSight region

Select a region.
US East (N. Virginia) ▼ **US East (N. Virginia)を選択**

QuickSight account name
Enter a unique QuickSight account name
You will need this for you and others to sign in. **任意のアカウント名(他の方と被らないもの)**

Notification email address
myemail@example.com **ご自身のメールアドレス**
For QuickSight to send important notifications.

Enable invitation by email
☒ Allow inviting new users by email. This setting cannot be changed after sign-up is complete.

☒ Enable autodiscovery of data and users in your Amazon Redshift, Amazon RDS, and AWS IAM services.

☐ Amazon Athena
Enables QuickSight access to Amazon Athena databases

☐ Amazon S3
Enables QuickSight to auto-discover your Amazon S3 buckets [Choose S3 buckets](#)

☐ Amazon S3 Storage Analytics
Enables QuickSight to visualize your S3 Storage Analytics data

☐ AWS IoT Analytics
Enables QuickSight to visualize your IoT Analytics data

[Finish](#)

この画面では、QuickSight アカウントの設定を行ないます。QuickSight capacity region は、QuickSight の Author ユーザに付属する SPICE 容量がどのリージョンに確保されるかを指定するものです。今回は US East(N. Virginia)を選択してください。なおここで選択した以外のリージョンにも追加で SPICE 容量を追加する事は可能です（追加 SPICE 容量の費用が必要です）。

QuickSight account name は任意の名前を付けます（アルファベット、数字、ハイフンのみ利用可能）。

Notification email address には、ご自身のメールアドレスを入れてください。

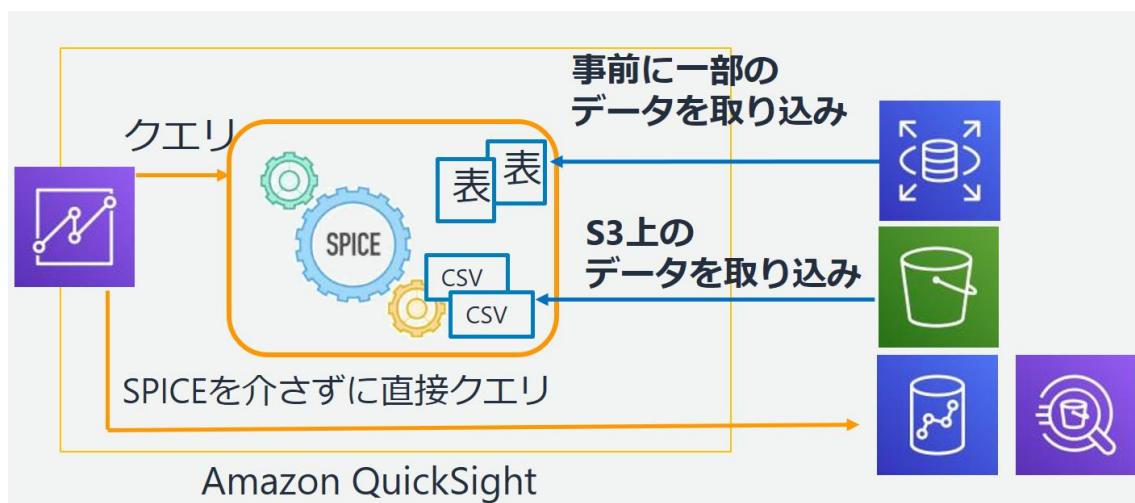
加えて上記のように”**Enable autodiscovery of...**”にチェックし、Finish を押してください。

これで QuickSight が利用可能になりました。

事前準備：CSV ファイルをアップロードし、データセットを作成する

QUICKSIGHT と SPICE

QuickSight には SPICE というインメモリ型的高速データベースが内蔵されています。この SPICE の領域にデータを取り込んでおくことで、高速な BI 応答が可能になります。もちろん、SPICE を使わず、Athena や Amazon Redshift に直接クエリ（SQL）を発行して利用することも可能です。



本ハンズオンでは、CSV ファイルを SPICE 領域にアップロードし、それを可視化します。

CSV ファイルを確認する

このハンズオンで使用する CSV ファイルは、このハンズオン資料に同梱されています。もしくはハンズオンの講師が指示する URL からダウンロードしてください。

CSV ファイルは 2 つあります。Patient-Info.csv ファイルをテキストエディタ等で開くと、以下のようにカンマ区切りでデータが格納されている事が分かります。最初の行は列名です。

Patient ID（患者 ID）から始まっており、Admit Date（入院日付）等の日付は、日/月/西暦（2 桁）の形式で記録されている事が分かります。（※この CSV を EXCEL で見ると、日付部分の解釈の違いにより、以下とは異なる見え方になるので注意が必要です）

```

1 Patient ID,Admit Date,Discharge Date,Priority,Hospital,
2 3,10/14/16,10/21/16,Low,North Hospital,-213.25,38.94,35
3 6,2/20/18,2/21/18,Not Specified,North Hospital,-4.64,2.
4 32,7/15/17,7/16/17,High,North Hospital,-128.38,8.46,8.9
5 32,7/15/17,7/16/17,High,South Hospital,-1748.56,70.89,8
6 32,7/15/17,7/17/17,High,North Hospital,-85.129,7.99,5.0
7 32,7/15/17,7/17/17,High,North Hospital,1054.82,107.53,5
8 35,10/22/17,10/23/17,Not Specified,North Hospital,60.72
9 35,10/22/17,10/24/17,Not Specified,North Hospital,48.98

```

もう一つの Assign.csv ファイルは小さいファイルです。テキストエディタ等で開くと各社員 (Account Rep) が担当している Region (地域) を表すマスター表であることが分かります。この Region 列は、Patient-Info.csv にも存在します。

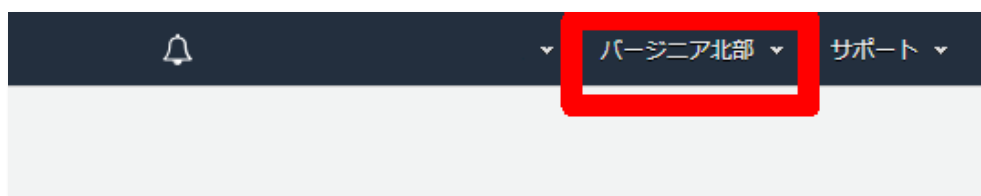
```

1 Account Rep,Region↓
2 Kevin,East↓
3 Jose,Central↓
4 Susan,South↓
5 Mark,West↓

```

SPICE に 1 つ目の CSV ファイルをアップロードする

(まだの場合は) AWS マネジメントコンソールにログインし、利用リージョンを QuickSight 登録時に QuickSight region に指定したリージョンに切り替えます。今回のハンズオンで初めて QuickSight を利用開始した場合はバージニア北部 (N. Virginia) を選択します。事前に QuickSight アカウント作成済であった場合は、そのアカウントの QuickSight region に切り替えてください (以前に QuickSight を東京リージョンで開始していた場合は東京を選択)。

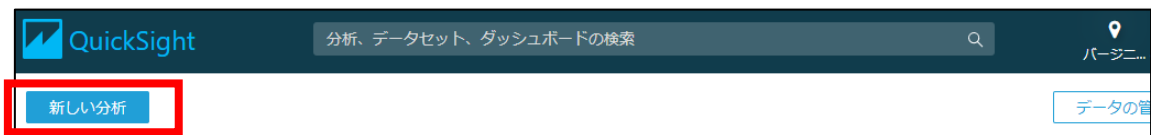


サービス一覧から QuickSight を選択します。

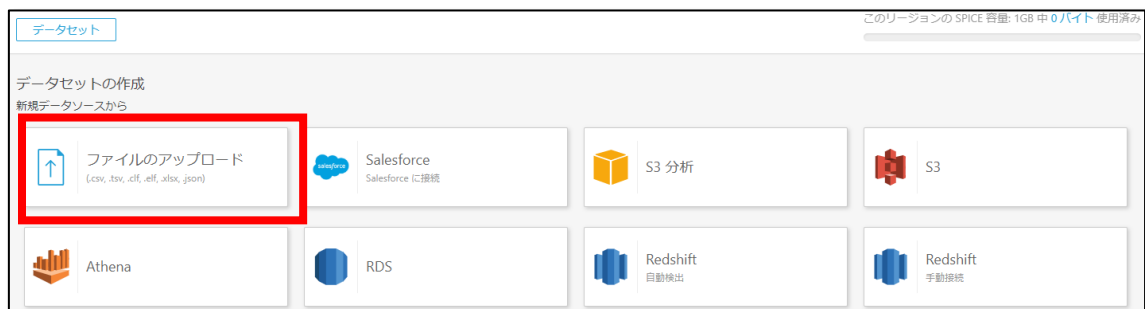


QuickSight の画面でも念の為、先程指定したリージョンが選択されていることを確認してください。（画面が英語表示の場合は右上の人型のアイコンから日本語に変更できます）

確認できたら、左上の”新しい分析”をクリックします。



次の画面で左上の”新しいデータセット”をクリックし、さらに”ファイルのアップロード”をクリックします。



ファイル選択のダイアログで Patient-Info.csv の方選択してアップロードします。確認画面が出ますので、そこで”設定の編集とデータの準備”を選択します。

ファイルのアップロード設定の確認

×

[詳細はこちら](#) to adjust the file upload settings.

設定

csv ファイル、 Patient-Info.csv

Patient ID	Admit Date	Discharge ...	Priority	Hospital	Profi
3	2016-10-14T...	2016-10-21T...	Low	North Hospital	-213.2
6	2018-02-20T...	2018-02-21T...	Not Specified	North Hospital	-4.64
32	2017-07-15T...	2017-07-16T...	High	North Hospital	-128.3
32	2017-07-15T...	2017-07-16T...	High	South Hospital	-1748
32	2017-07-15T...	2017-07-17T...	High	North Hospital	-85.12

設定の編集とデータの準備

次へ

すると「準備（PREPARE）」の GUI に移動します。

データセットを準備（PREPARE）する

以下の画面は「準備（PREPARE）」ツールで、ここでは分析の準備として、データを利用しやすい形にすることができます。例えばフィールド（列）の名前を変えたり、不要なフィールドを見えなくしたり、事前に複数のデータソースをジョイン（結合）する等です。

データソース: SPICE

フィールド: すべてのフィールドを選択済み

計算フィールドを追加

Augment with SageMaker

フィールドの検索

計算フィールド: 計算フィールドがありません。

選択: すべて | なし

選択: ☒ Patient ID, ☒ Admit Date, ☒ Discharge Date, ☒ Priority, ☒ Hospital, ☒ Profit, ☒ Price, ☒ Cost, ☒ Revenue, ☒ Discount

データ: Patient-Info.csv

Patient ID	Admit Date	Discharge D...	Prior...	Hosp...	Profit	Price	Cost	Reve...	Disc...	St
3	2016-10-14T00:00:00...	2016-10-21T00:00:00...	Low	North Hospital	-213.25	38.94	35.0	326.925	0.04	Persc
6	2018-02-20T00:00:00...	2018-02-21T00:00:00...	Not Specified	North Hospital	-4.64	2.08	2.56	8.6625	0.01	Medi
32	2017-07-15T00:00:00...	2017-07-16T00:00:00...	High	North Hospital	-128.38	8.46	8.99	175.7	0.04	Medi
32	2017-07-15T00:00:00...	2017-07-16T00:00:00...	High	South Hospital	-1748.56	70.89	89.3	2201.75	0.09	Medi
32	2017-07-15T00:00:00...	2017-07-17T00:00:00...	High	North Hospital	-85.129	7.99	5.03	200.2919	0.04	Medi

本ハンズオンでは、以下 3 点の調整を行ないます。

1) 他データソースのジョイン

もう一つの CSV データをここで追加し、Region フィールド（列）を基準にジョインします。画面上方の”データを追加”をクリックし、ダイアログから”ファイルのアップロード”を選択します。ファイルは、先程ダウンロードした中にある Assign.csv を選択します。

データを追加 ×

データソース

Patient-Info.csv CSV

データソースの切り替え

テーブル

☐ Patient-Info.csv

ファイルのアップロード 選択

ファイルのアップロード設定の確認の画面では”次へ”を押します。

ファイルのアップロード設定の確認 ×

[詳細はこちら](#) to adjust the file upload settings.

設定

csv ファイル、Assign.csv

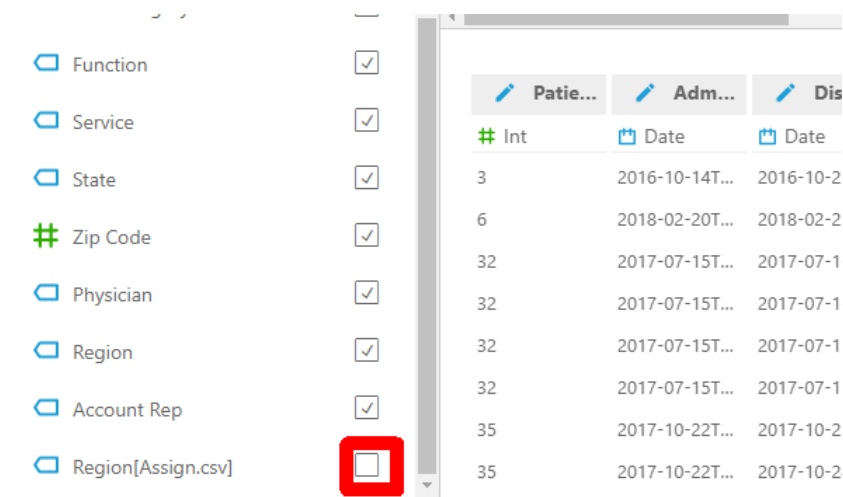
Account Rep	Region
Kevin	East
Jose	Central
Susan	South
Mark	West

次へ

アップロードした後は2つのデータの結合方法を指定します。2つの円のようなアイコン（上部）をクリックし、画面下に出る結合設定で結合（ジョイン）する列を指定します。今回は Region フィールドでジョインしたいので、両方とも Region を指定し、結合タイプは Inner のままにして、”適用”を押します。



これで、画面左側のフィールドに、Assign.csv にあった Account Rep と Region[Assign.csv]が追加されます。Region 列は Patient-Info にも存在し、2つある必要はないため、Region[Assign.csv]はチェックボックスを外して表示しないようにします。

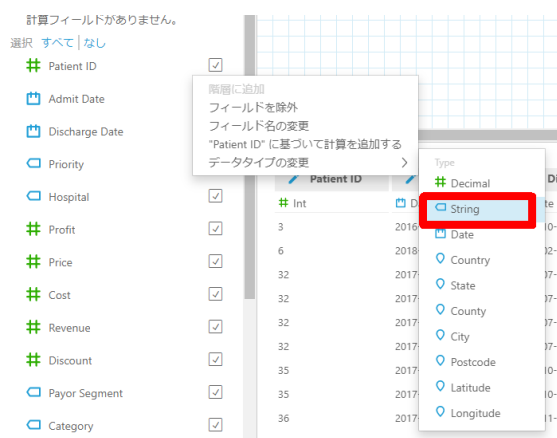


今回のハンズオンでは CSV ファイルと CSV ファイルをジョインしましたが、異なるデータソース間、例えば PostgreSQL データベース上の表と、MySQL データベース上の表と、CSV ファイルをジョインする事も可能です。

(補足：異なるデータソースをまたがってジョインされたデータは、SPICE 内に保存されません。SPICE の容量を消費する点に注意が必要です。)

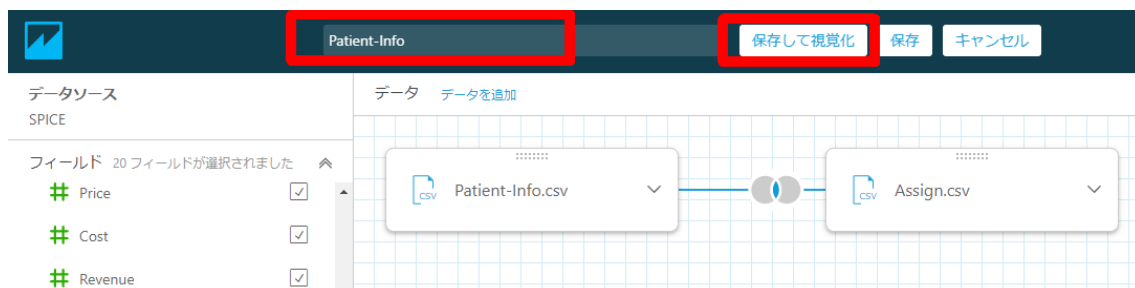
2) Patient-ID フィールドの型を修正

Patient-ID フィールドが自動的に Int (整数) 型に認識されていますが、正しくは文字列なので、String 型に変更します。左側のペインから Patient ID にカーソルをあわせると"V"が表示されるので、クリック=>データタイプの変更=>String を選択します

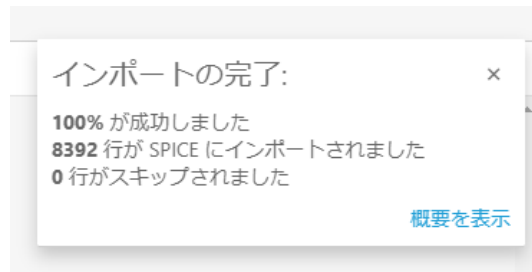


3) 名前の変更と保存

画面最上部にある枠で、このデータセットの名前をつけることが可能です。任意の文字列に変更してください。ここではデフォルトの「Patient-Info」のままとしました。



最後に、“保存して視覚化”を押し、分析画面に移動します。SPICE にデータが取り込まれるまで少し待ちます。右上の表示が“インポートの完了”になったらデータの取り込みが完了しています。（このダイアログは x を押して閉じてください）





データ取り込みが完了すると上記のような画面が表示されています。これは分析（Analysis）を作っていく画面です。左側にフィールド（列）と、ビジュアル（グラフ）の一覧が表示されています。まずビジュアルタイプ（グラフの種類）を選択し、それに必要なフィールドを選択することで可視化されます。ビジュアルは左上の「+」を押すことで複数追加できますし、サイズや位置を変更することが可能です。

操作は自動的にセーブされます。また、画面左上にある「元に戻す」ボタンで操作を取り消す（UNDO する）事が可能です。

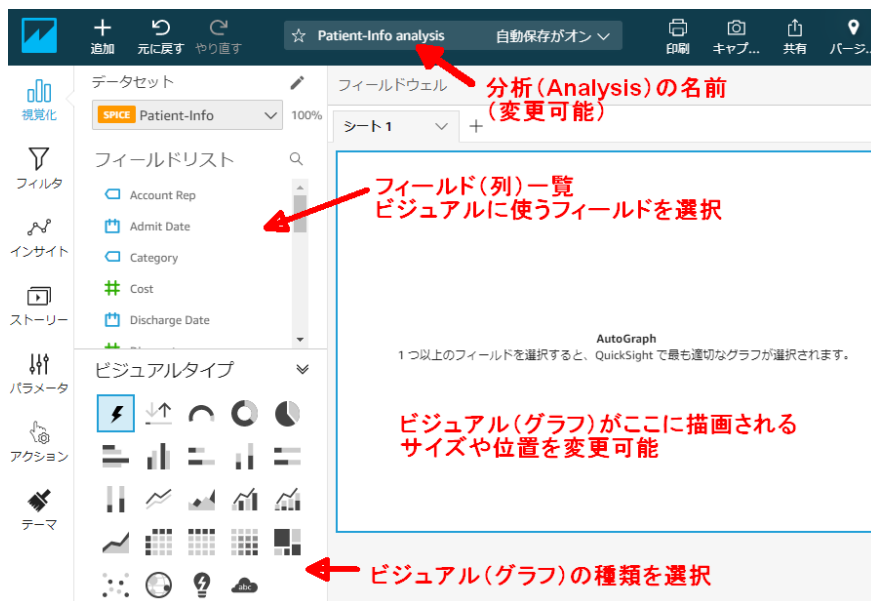
ここからはハンズオン 1 を実施しますが、ハンズオン 1 と 2 の内容は独立しているため、セルフラーニングで実施されている方はハンズオン 1 を飛ばしてハンズオン 2 を実施していただいても問題ありません。

ハンズオン 1：基本的な可視化とクイックフィルタ

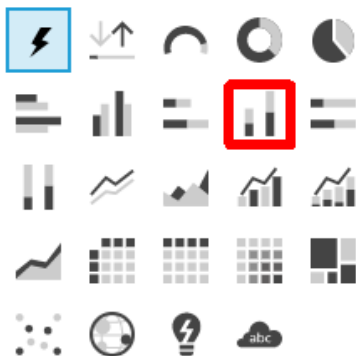
ハンズオンを開始するには前章の「事前準備」を完了している必要があります。

垂直積み上げ棒グラフの作成

ここからは SPICE に取り込んだデータを可視化していきます。




ここまでの準備で、分析(Analysis)作成の画面（上記）になっており、1つ空白のビジュアル（グラフ）があります。デフォルトではビジュアルタイプは”AutoGraph”になっています。これは選択したフィールドによって自動的にビジュアルタイプを選択する機能ですが、今回は垂直積み上げ棒グラフを選択します。（上から2段目、左から4番目のアイコン）



選択すると、右側のビジュアルの上に”フィールドウェル”が準備されます。

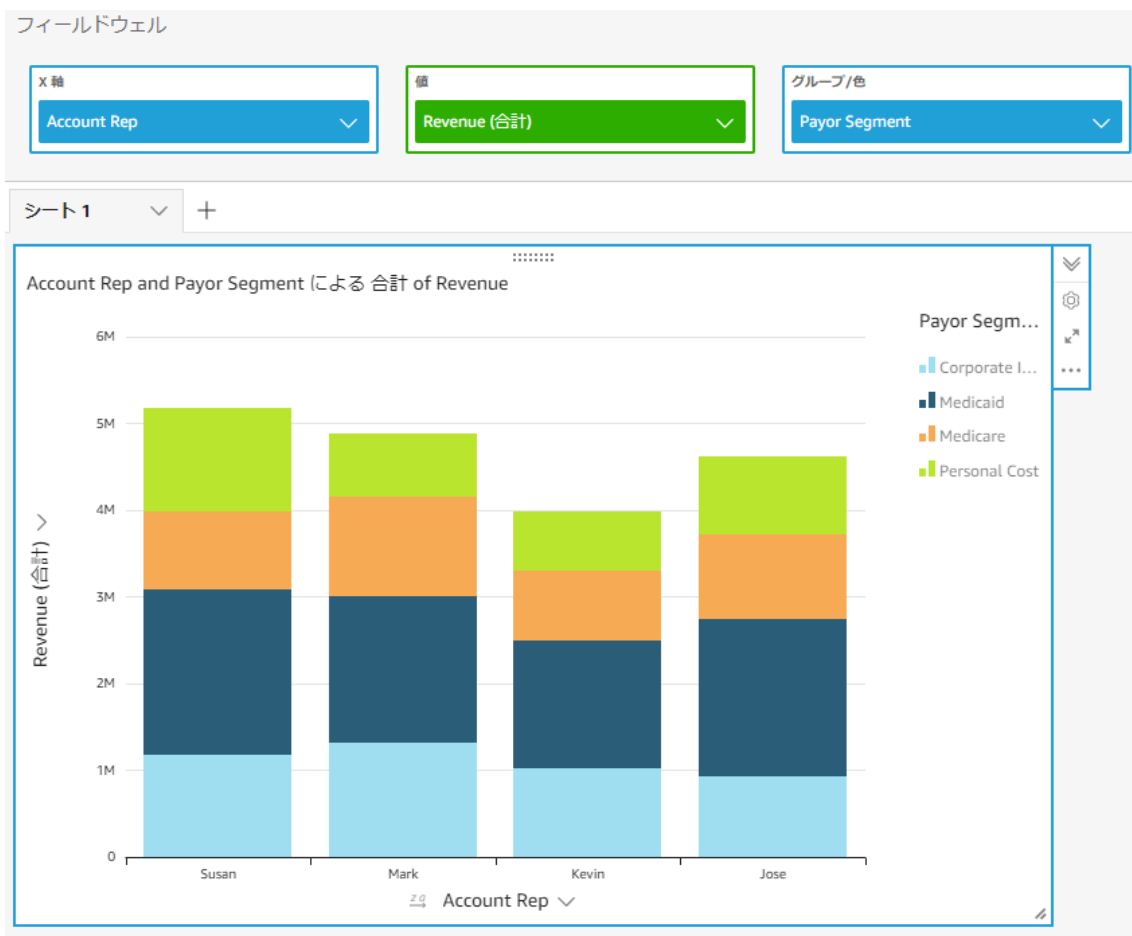


※もしフィールドウェルが上記のように表示されない場合は、右上にある  マークをクリックして、フィールドウェルを開いてください

この3つの領域に、左からフィールドを Drag & Drop (以下、D&D) することでそのフィールドのデータをもとに可視化が行われます。まず X 軸（横軸）から設定します。担当者(Account Rep)ごとの売上(Revenue)を見たいため、X 軸には Account Rep を D&D します。




同様に値には、”Revenue”（売上）を、グループ/色には”Payor Segment”（支払いセグメント）を D&D します。

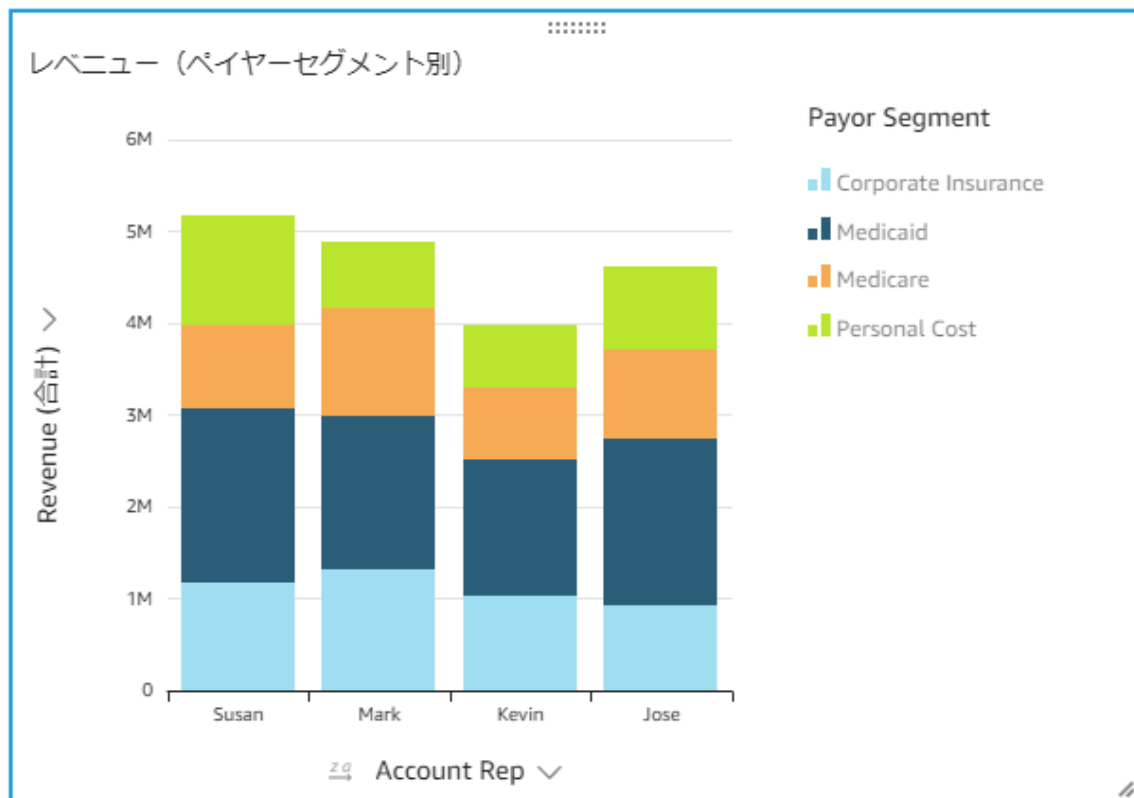


上記のようなビジュアルが出力されました。このビジュアルは担当者ごとの Revenue を合計したものを”Payor Segment”別に積み上げて表示しています。Revenue が合計されているのは、上部フィールドウェルの”値”が”Revenue(合計)”になっているためです、この集計方法を変えた場合は、その横にある”V”をクリックする事で”平均”や、”カウント”（個数）等に変えることができます。（今回は合計のままで進めます）

ビジュアル上部にある”Account Rep and Payor Segment による 合計…”といった文字列は自動的に作成されたこのビジュアルのタイトルですが、変更可能です。文字列をクリックして、分かりやすい表現、例えば「レベニュー（ペイヤーセグメント別）」としてください。

また、ビジュアル左側にあるレジェンド（凡例）の領域が小さくて文字が見切れています。ビジュアル上にマウスカーソルを置くと、のようなマークが表示されるので、そこをドラッグしてサイズを変えることができます。これで1つ目のビジュアルが完成しました。（下図）

このビジュアルで使っている Account Rep フィールドは、Assign.csv から、他は Patient-Info.csv から取得しています。手軽に複数のデータをジョインして可視化する事ができるのが分かります。

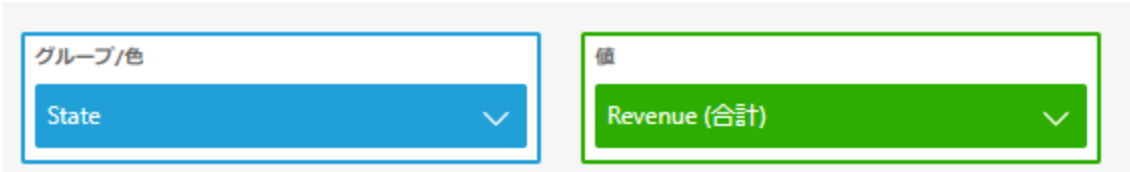


ドーナツグラフの作成

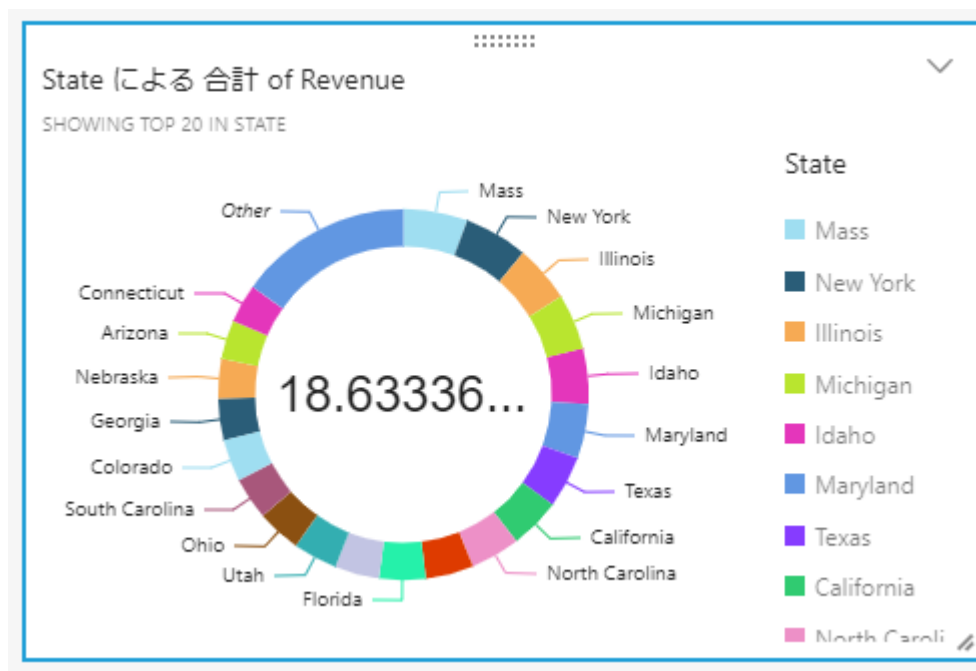
次はドーナツグラフを作成して、レベニュー（売上）が、どの州でどれぐらいの比率になっているかを可視化します。

左上の”+追加”->“ビジュアルを追加”を押すと新しいビジュアルが追加されます。その新しいビジュアルが選択された状態で、ビジュアルタイプとして”ドーナツグラフ”（最上段、左から4つ目）を選択してください。

フィールドウェルには、”グループ/色”と、”値”が用意されます。州ごとの売上比率を見たいので”グループ/色”には State（州）を、”値”には Revenue を D&D してください。




以下のようなビジュアルが表示されます。



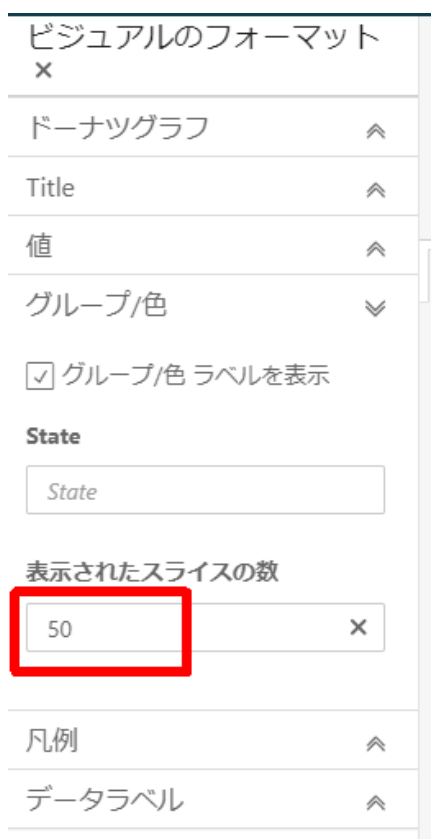
ドーナツ上になっている円の中心にあるのが、Revenue の合計です。その周辺を州ごとの売上合計が表示されていますが、ビジュアル左上に薄い字で”SHOWING TOP 20 IN STATE”と書

かれているのが分かります。これは自動的に上位 20 までに絞った表示にしているという意味です。トップ 20 に入らなかった州は”Other”に積み上げられています。

あまり表示増やしても見づらくなりますが、今回はあえて全て表示させてみましょう。

ビジュアル右側にあるツールバーからギアのマーク () をクリックすると、左側に”ビジュアルのフォーマット”ペインが表示されます。

その中から”グループ/色”を開き、”表示されたスライスの数”に 50 と入れて ENTER キーを押してください。

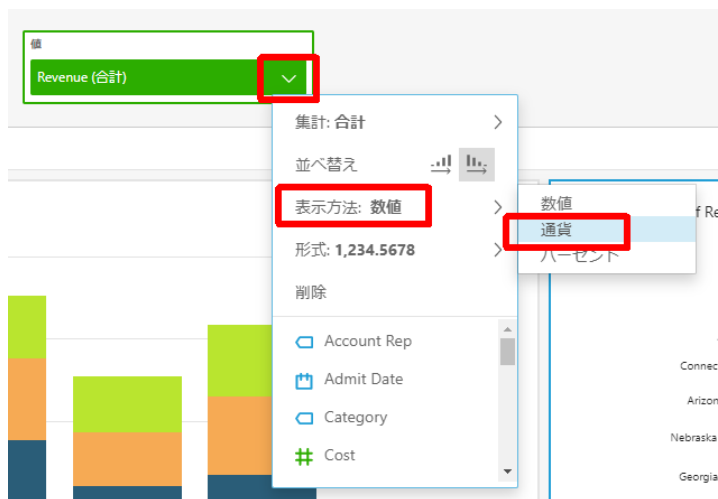


米国は 50 州なので、これで全てのデータが表示されます。ただし、ブラウザのサイズによっては、州の名前（ラベル）が表示しきれない場合があります。その場合はブラウザサイズを広げたり、適切なスライス数、例えば 15 に変えたりしてください。（スライス数をいくつにし

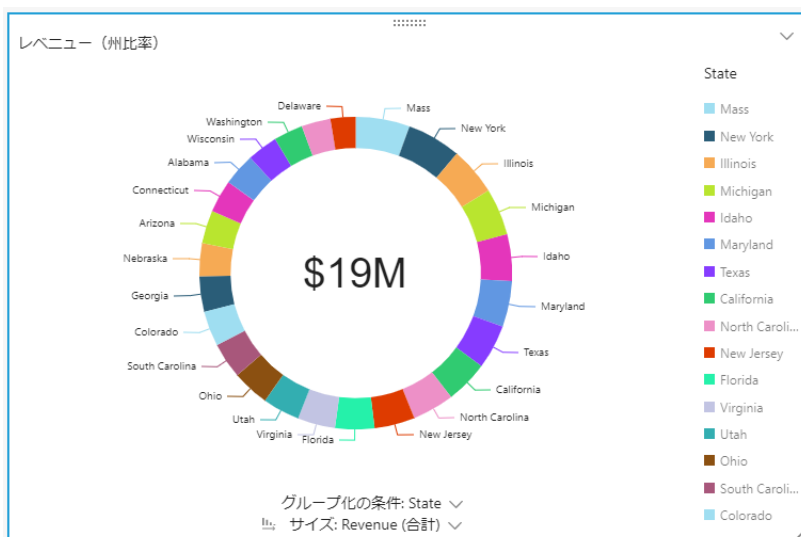
ても、この後のハンズオンには影響がありません）。[https://us-east-](https://us-east-1.quicksight.aws.amazon.com/sn/analyses/b947dead-9326-460c-8ffb-1933de1caf12)

[1.quicksight.aws.amazon.com/sn/analyses/b947dead-9326-460c-8ffb-1933de1caf12](https://us-east-1.quicksight.aws.amazon.com/sn/analyses/b947dead-9326-460c-8ffb-1933de1caf12)

次に円中央の値の表示を変更します。18.6336… のような数字が表示されていますが、Revenue は通貨（ドル）なので、分かりやすい形に変更します。フィールドウェルの”値”でVをクリックし、表示方法=>通貨を選択してください。



最初のビジュアルと同様に、ビジュアルのタイトルを”レベニュー（州比率）”に変更します。これで、以下のようなビジュアルが完成しました。



ピボットテーブルとコンディショナルフォーマットの作成

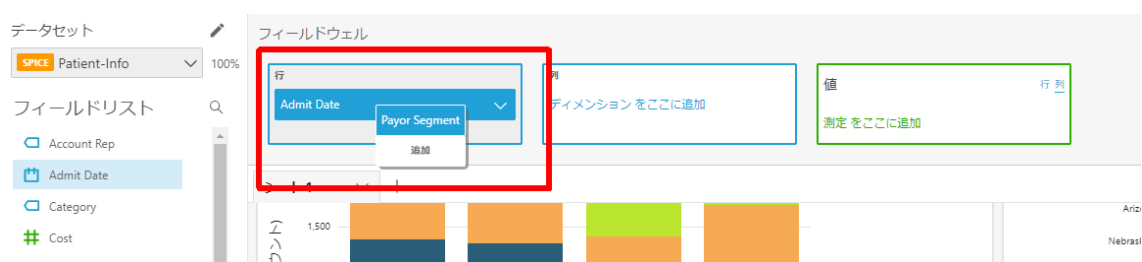
次はピボットテーブルを使って、レベニュー（売上）の詳細データを表示し、それにコンディショナルフォーマットを追加することで、売上が良いところは緑色、悪いところには赤色を表示させます。

左上の”+追加”->“ビジュアルを追加”を押すと新しいビジュアルが追加されます。その新しいビジュアルが選択された状態で、ビジュアルタイプとして、“ピボットテーブル”（上から4段目段、左から2つ目）を選択してください。

フィールドウェルには、“行”、“列”、“値”が用意されます。

行にも列にも、複数のフィールドを指定して階層化できるのがピボットテーブルの特徴です。フィールドリストの Admit Date（入院日）をクリックしてください。（D&D 無しで）自動的にフィールドウェルの”行”に Admit Date が挿入されます。これはフィールドの型や定義から自動的に判断してフィールドが追加される仕組みです。続いて、Payor Segment をクリックすると、Admit Date の下に配置されます。これで階層化が実現できます。

もしくは、Admit Date の下あたりに Payor Segment を D&D する方法でもかまいません。以下の図のように、“追加”と表示される位置までドラッグしてからドロップしてください。



以下のようにフィールドウェルの”行”で2つのフィールドが階層化されているはずです。

データセット: SPICE Patient-Info 100%

フィールドリスト

- Admit Date
- Category
- Cost
- Discharge Date
- Discount
- Function
- Hospital
- Patient ID
- Payor Segment
- Physician
- Price

フィールドウェル

行: Admit Date, Payor Segment

列: デイメンションをここに追加

値: 測定をここに追加

シート 1

Admit Date and Payor Segment によるレコードのカウント

Admit Date	Payor Segment	Count
1月 1, 2015 1...	Corporate Insura...	2
	Medicaid	1
	Medicare	7
1月 2, 2015 1...	Corporate Insura...	1
	Personal Cost	1
1月 3, 2015 1...	Medicaid	3
	Medicare	3

このような方法を使って、以下の図に合わせて、行と値のフィールドウェルを設定してください。行は階層を表しているため、上下の並びが下図を同じようになるようにしてください。

また、必要に応じて見やすいサイズまでピボットテーブルの横幅を広げてみてください。（ビジュアルの右下をドラッグするとリサイズできます）

フィールドウェル

行: Admit Date, Payor Segment, Hospital, Category

列: デイメンションをここに追加

値: Revenue (合計), Profit (合計), Cost (合計)

シート 1

Admit Date, Payor Segment, Hospital, and Category による 合計 of Revenue, 合計 of Profit, and 合計 of Cost

Admit Date	Payor Segment	Hospital	Category	Revenue	Profit	Cost
1月 1, 2015 1...	Corporate Insura...	North Hospital	Step Down	906.7875	245.97	5.76
			Surgical	235.9125	-32.48	5.83
		Medicaid	South Hospital	Ante/Post Partum	1,555.2375	-131.31
	Medicare	North Hospital	BMT	97.6	1.74	0.75
			Critical Care	2,526.4338	-60.39	5.99
			Medical	3,437.6338	667.449	5.26
			Psychologych	300.375	-154.6175	5.68
			Surgical	16,001.875	-363.7	36.99
	South Hospital	Critical Care	816.025	-1,358.0	52.02	

ピボットテーブルが構成されました。”-“マークをクリックすると各セルを閉じたり開いたりすることが可能です。

次に日付表示をカスタマイズします。フィールドウエルの”行”にある”Admit Date“のVをクリックし、形式=>その他フォーマットオプションを選択します。

左に出るデータのフォーマットペインで、”日付”を選択し、好みのフォーマット、例えば”2020年9月20日”等を選択します。

また、フィールドウエルの”値”にある3つのフィールド(Revenue、Profit、Cost)それぞれについて、Vをクリックし、”表示方法”=>”通貨”を選択して”\$”付きの表記に変更してください。これで、以下のような表示になります。

Admit Date, Payor Segment, Hospital, and Category による 合計 of Revenue, 合計 of Profit, and 合計 of Cost						
Admit Date	Payor Segment	Hospital	Category	Revenue	Profit	Cost
2015年1月1日	Corporate Ins...	North Hospital	Step Down	\$906.79	\$245.97	\$5.76
			Surgical	\$235.91	-\$32.48	\$5.83
	Medicaid	South Hospital	Ante/Post Partum	\$1,555.24	-\$131.31	\$43.32
	Medicare	North Hospital	BMT	\$97.60	\$1.74	\$0.75
			Critical Care	\$2,526.43	-\$60.39	\$5.99

最後に、コンディショナルフォーマットを設定します。これは条件に合わせて表示（色やアイコン）を変化させる機能です。ここでは、Profit（利益）が、マイナスである場合に背景が赤色に、\$100 以上の場合には緑色になるように設定します。

ビジュアル右端ツールバーの”...”から、”条件付き書式設定”をクリックします。

新しい背景色

列
Profit

塗りつぶしタイプ
☒ 固体 ☐ 勾配

に基づくフォーマットフィールド
Profit

集計
合計

条件 - 値
次より小さい

値
0

色
色

+ 条件を追加する

列の選択では色変更の対象とする”Profit”を選択し、次に”背景色を追加する”を選択します。

フォーマットフィールドでは、”条件-値”を”次より小さい”にし、値は0にします。その横にある、色の指定では赤色を指定します。これでマイナスの値のときに背景が赤色になります。

続いて、値を指定した部分の下にある”+条件を追加する”をクリックします。

すると、追加の条件が指定できるようになるため、そこでは”条件#2 - 値”に、”次以上”を指定し、値には 100 を、色は緑色を指定し、最後に”適用”を押します。

これで以下のようなピボットテーブルが実現できました。タイトルここでは”レベニュー/プロフィット/コスト表”に変更しています。

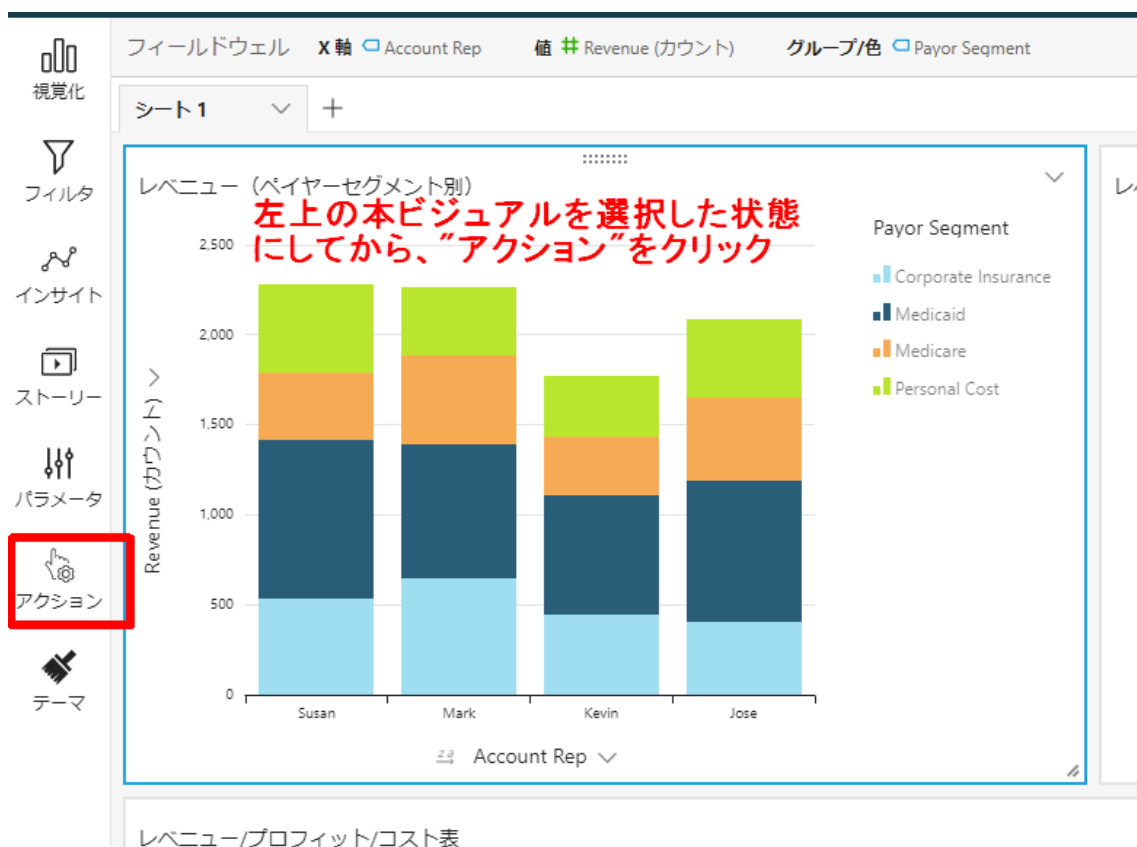
レベニュー/プロフィット/コスト表						
Admit Date	Payor Segment	Hospital	Category	Revenue	Profit	Cost
☐ 2015年1月1日	☐ Corporate Ins...	☐ North Hospital	Step Down	\$906.79	\$245.97	\$5.76
			Surgical	\$235.91	-\$32.48	\$5.83
	☐ Medicaid	☐ South Hospital	Ante/Post Partum	\$1,555.24	-\$131.31	\$43.32
	☐ Medicare	☐ North Hospital	BMT	\$97.60	\$1.74	\$0.75
			Critical Care	\$2,526.43	-\$60.39	\$5.99
			Medical	\$3,437.63	\$667.45	\$5.26
			Psychologych	\$300.38	-\$154.62	\$5.68
			Surgical	\$16,001.88	-\$363.70	\$36.99
		☐ South Hospital	Critical Care	\$816.93	-\$1,358.90	\$53.03
☐ 2015年1月2日	☐ Corporate Ins...	☐ Downtown H...	Surgical	\$1,090.60	-\$342.91	\$35.00
	☐ Personal Cost	☐ North Hospital	Medical	\$225.45	-\$111.80	\$4.69

クイックフィルタ機能による絞り込みの実現

次はクイックフィルタを分析全体に適用します。フィルタとは、条件を付けてデータを絞り込む機能のことで、自分が見たいデータ範囲だけを表示させるために利用します。

QuickSight でフィルタを作成する方法は複数ありますが、ここでは手軽に作ることができるクイックフィルタを作成します。クイックフィルタは、あるビジュアルの上でクリックした際に、そのクリックされた値に応じて他ビジュアルがフィルタされる（絞り込まれる）機能です。

今回は一番初めに作った左上のビジュアルへのクリックをトリガーにフィルタしたいので、まず左上の「レベニュー（ペイヤーセグメント別）」ビジュアルを選択した状態にしてから（選択するとそのビジュアルは青い線で囲まれます）、画面左端にある「アクション」のアイコンをクリックします。



表示された”アクション”ペインで、”クイック作成”の下にある”同じシートのビジュアルをフィルタ”を選択します。すると、”アクション 1”が作成されるので、その右側にある”V”=>編集をクリックします。

図のようなアクションの設定ペインが表示されます。

”アクティベーション”は、選択（クリックするとすぐフィルタされる）と、メニューオプション（一旦メニューが出る）から選択できます。今回はデフォルトの”選択”のまま使用します。

アクションタイプはフィルタアクション（フィルタする）と URL アクション（別の URL に遷移する）が選択できます。こちらもデフォルトのフィルタアクションのままにします。

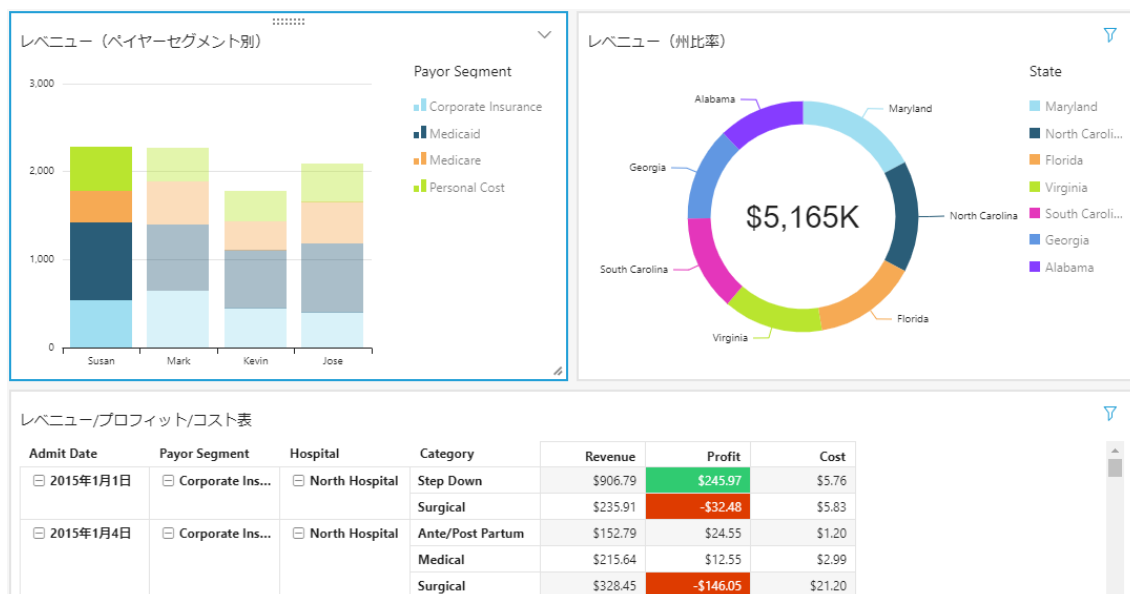
フィルタ範囲は、クリックされた際、どのフィールドの値を元にフィルタするかを選択します。今回は”選択されたフィールド”を指定し、”Account Rep”だけにチェックを入れます。これにより、グラフをクリックした際に、その Account Rep だけの情報になるようフィルタされます。

もしデフォルトの”すべてのフィールド”を選択した場合は、クリックした位置の Account Rep & Payor Segment 両方の条件を満たすデータだけに絞り込まれます。

最後のターゲットビジュアルは、このフィルタがどのビジュアルに反映されるかを指定します。今回はデフォルトであるすべてのビジュアルのままにし、”保存”を押します。

これでクイックフィルタは完成です。動作確認のために、左上の”レベニュー（ペイヤーセグメント別）”の棒グラフから、Susan の棒グラフをクリックしましょう。すると、Susan 以外の棒

グラフは半透明のような色に替わり、ドーナツグラフの数字や、ピボットテーブルの数字にも変化があるのが確認できるでしょう。



つまり、クリックした部分の Account Rep (Susan) でクイックフィルタが起動し、ドーナツグラフやピボットテーブルが Susan のデータだけになるよう絞り込まれた状態になっています。再度 Susan の棒グラフをクリックすると、フィルタが解除されます。他の人でもフィルタが効くか確認してください。

これでハンズオン 1 は完了です。時間が余った方は以下のオプション演習を実施していただくか、ハンズオン 2 に進んでください。

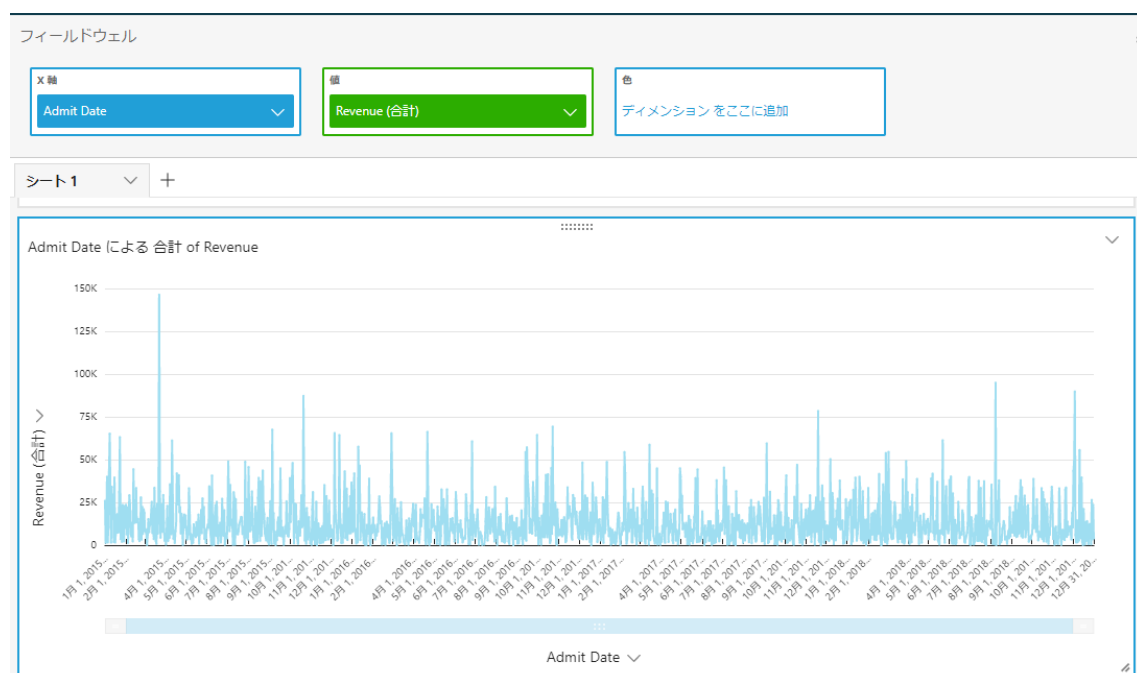
(オプション演習) 折れ線グラフと予測機能 (ENTERPRISE EDITION が必要)

レベニュー (売上) の折れ線グラフを作成し、QuickSight の ML インサイト機能で6ヶ月先の売上の予想を作成します。

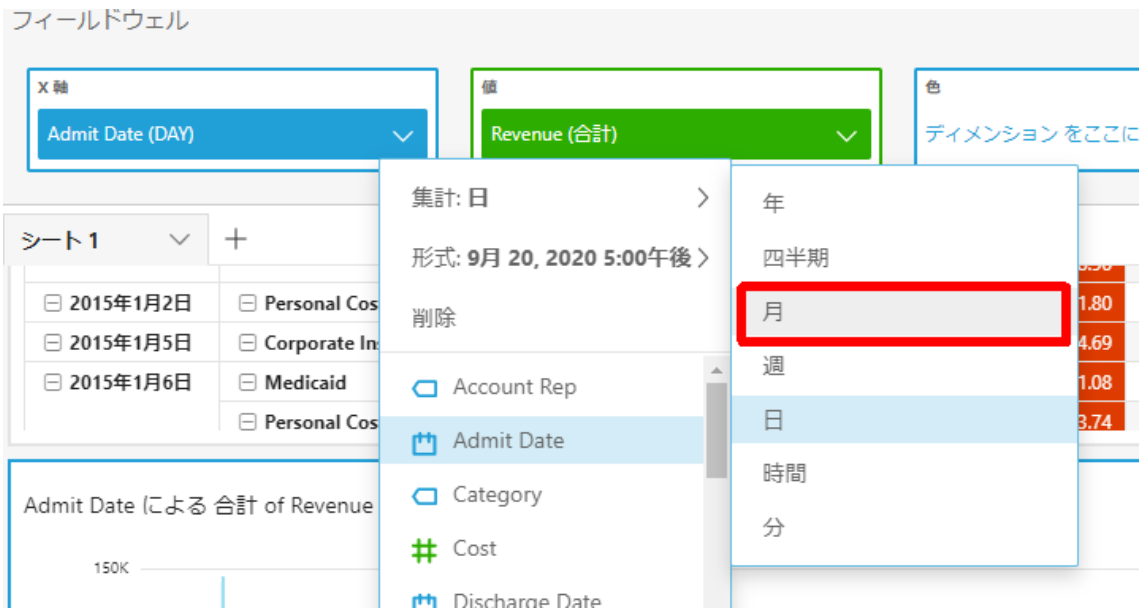
左上の"+追加"->"ビジュアルを追加"を押すと新しいビジュアルが追加されます。その新しいビジュアルが選択された状態で、ビジュアルタイプとして、"折れ線グラフ" (上から3段目、左から2つ目) を選択してください。また、ビジュアルの右下隅をドラッグして、横幅を大きくしてください。

フィールドウェルには、"X軸"、"値"、"色"が用意されます。今回は日時の経過によるレベニュー推移を表示するので、"X軸"には Admit date を、"値"には Revenue を D&D します。

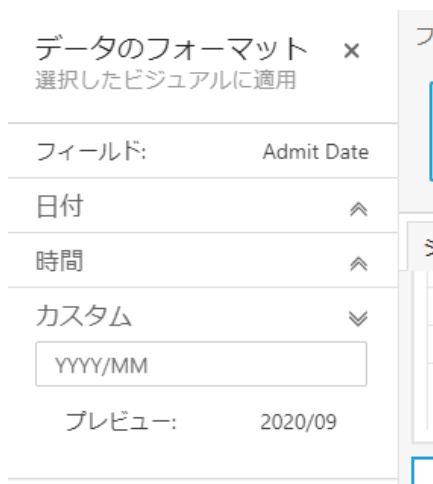
(D&D せず、クリックするだけでも適切なフィールドウェルに設定されます)



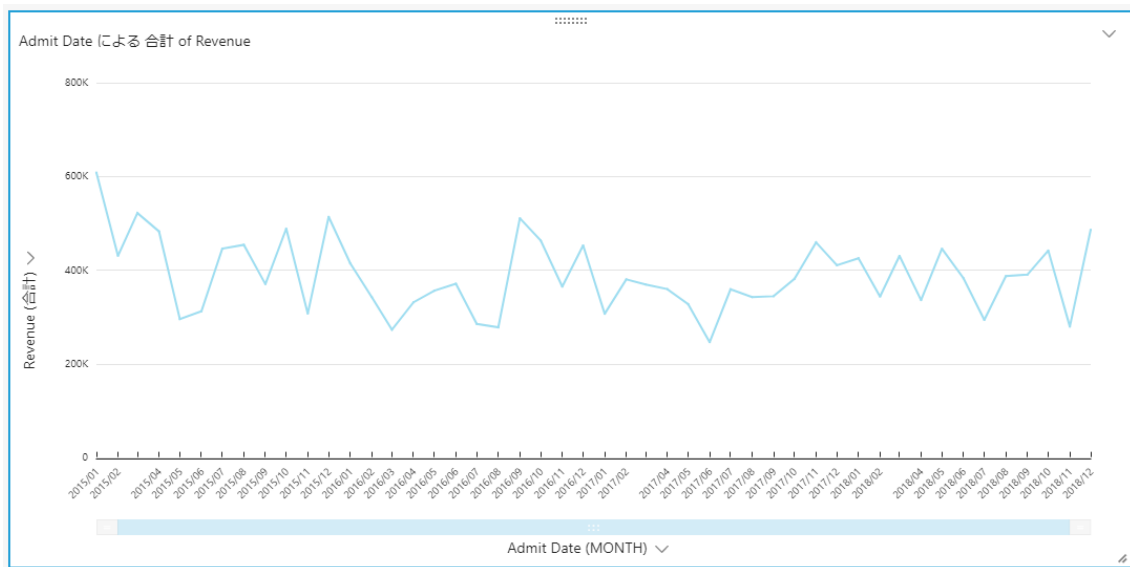
図のように、日付ごとのレベニューが表示されますが、粒度が細かすぎて見づらいため、月ごとのサマリに変更します。フィールドウェルの X 軸にある Admit Date の"V"をクリックして、集計=>月 を選択します。



同様に、Admit Date の”V”をクリックして、形式=>その他のフォーマットオプションを選択します。左ペインで、”日付”ではカスタムを選択、その下の方にある”カスタム”で YYYY/MM と入力します。



これで、以下のような表示になります。



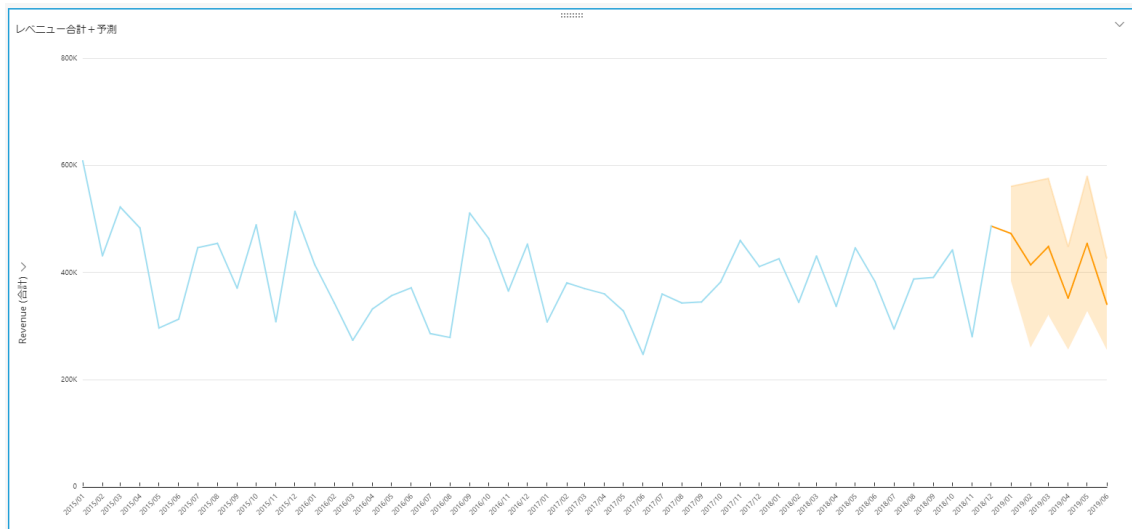
これで、レベニューの月ごとの推移が表示できましたので、これに予測を追加します。予測は QuickSight の ML（機械学習）インサイトの機能の 1 つで、与えられたデータから予測値を出力します。

この予測に使うモデルは、時系列に変化するデータに適した学習済のもので、ユーザ側では調整できません。しかし、（2020 年 4 月現在）プレビューで提供されている Amazon SageMaker との連携機能を利用すると、ユーザ独自に学習させたモデルを使った分析が可能です（本ハンズオンでは SageMaker との連携は扱いません）。

ではビジュアルの右上にある“…”から”予測を追加”をクリックします。

左ペインに予測プロパティが表示されます。ここで”期間を進める”が 14 になっていますが、今回は半年先までの予測がほしいのでここを 6 にし、適用をクリックします。

最後にビジュアルのタイトルを”レベニュー合計+予測”として、以下のようなビジュアルが完成しました。



上図で、青色の線で表示されているのが、実際の値、オレンジ色の線で表示されているのが先程設定した予測値です。あくまで機械的な予測ですので、絶対ではありません。また薄いオレンジ色で塗られた領域は、予測の「幅」を示しています。

(オプション演習) テーマを切り替える

QuickSightにはテーマ機能があります。これは画面の見た目を指定しておいたパレット（色の設定）に合わせて変えるものです。デフォルトでいくつかプリセットパレットが用意されているので、そちらに変更してみましょう。

画面左端の、“テーマ”アイコンをクリックして、“Starter themes”の下にある”Midnight”と書かれたテーマの”・・・”ボタン=>”適用”を押してください。画面の色合いやビジュアル間のボーダー等が変更されるのが確認できます。テーマの別名で保存を選択すると独自のテーマを作成可能になりますので、試してみてください。



ハンズオン 2：より進んだ可視化

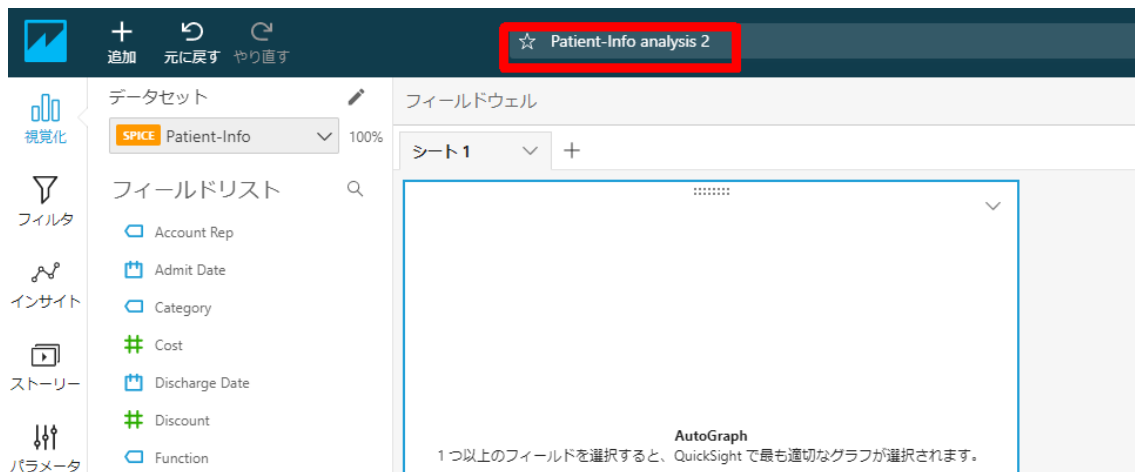
ハンズオン 2 を始める前に

ハンズオン 2 でも、ハンズオン 1 と同じデータセットを使用します。まだデータセットを作成していない方は、「事前準備」の章を確認して準備をしてください。

データセットが準備できたら、新しい分析（Analysis）を作成します。QuickSight トップ画面から（トップ画面に戻るには、画面左上の QuickSight アイコンをクリックします）、左上の”新しい分析”をクリックし、「事前準備」の章で作成したデータセット（Patient-Info）をクリック=>分析の作成をクリックし、新規の分析を作成します。



新しい分析では、一番上にある分析のタイトルを任意の名前に変更してください（デフォルトのままだと、ハンズオン 1 で作成したものと同名前で区別が付きにくくなるため）。ここでは”Patient-Info analysis 2”としました。



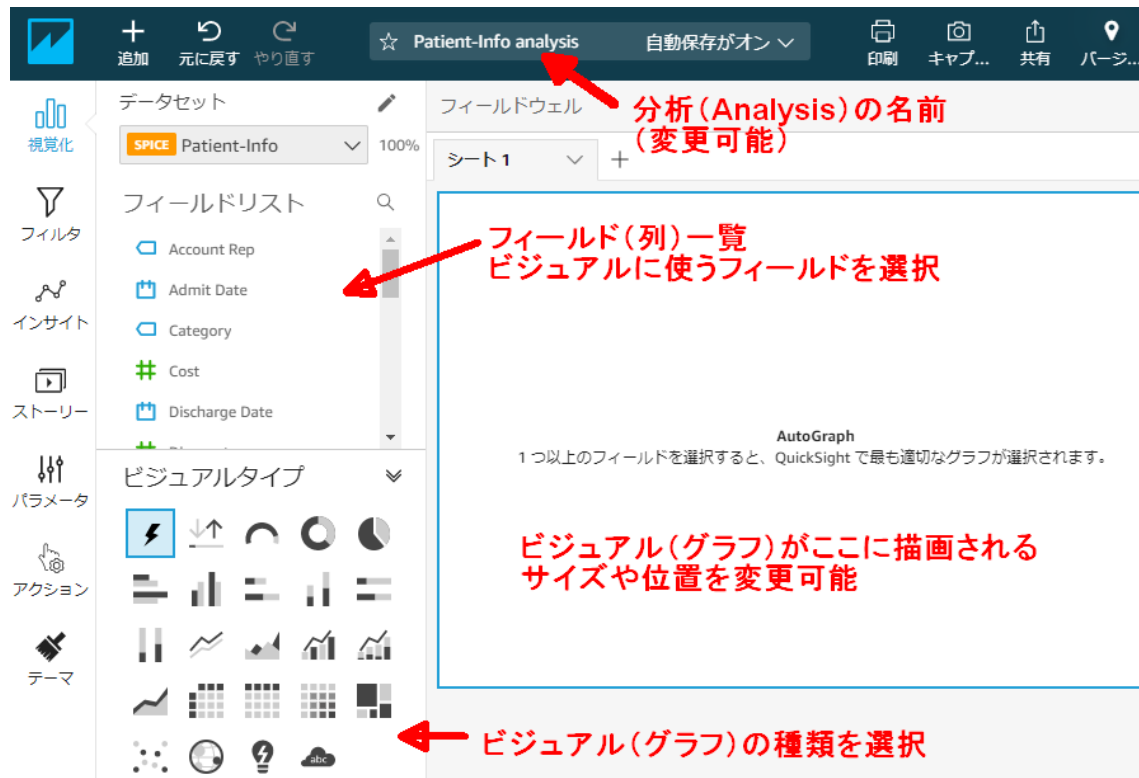
右上に“インポートの完了”ダイアログが出ている場合は x を押して閉じてください。



では、ハンズオン 2 を開始しましょう。

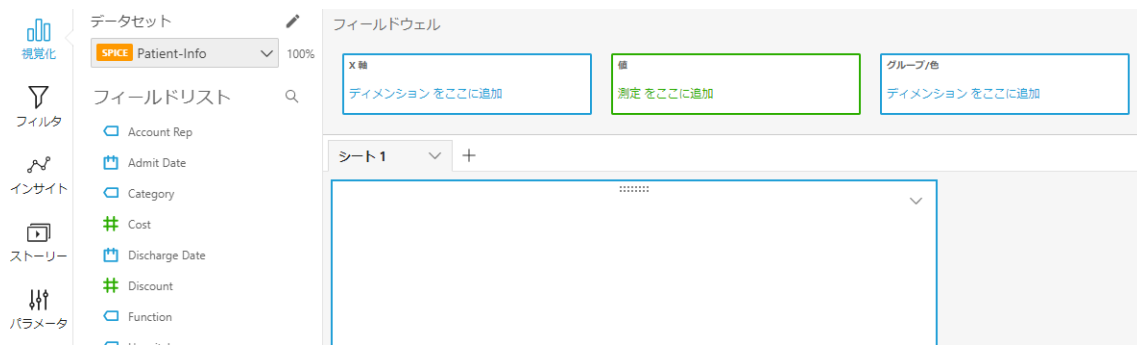
垂直棒グラフの作成とドリルダウンの実現

新規にビジュアルを作成すると以下のような画面が表示されます。



これは分析 (Analysis) を作っていく画面です。左側にフィールド (列) と、ビジュアル (グラフ) の一覧が表示されています。まずビジュアルタイプ (グラフの種類) を選択し、それに必要なフィールドを選択することで可視化されます。ビジュアルは左上の「+」を押すことで複数追加できますし、サイズや位置を変更することが可能です。

画面に空白のビジュアルが1つ表示されていますので、まずは左下のビジュアルタイプ一覧から”垂直棒グラフ” (上から2段目、左から2つ目) を選択してください。すると以下のような画面になります。



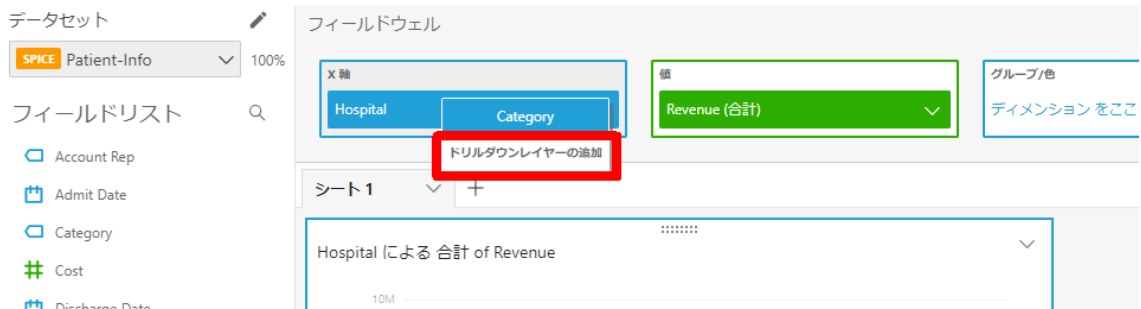
QuickSight での可視化の基本はビジュアルを追加した後に、フィールドをフィールドウェル上に設定していくことです。

※上図のようにフィールドウェルがビジュアルの上部に表示されない場合は、まだビジュアルタイプを選択していないか、もしくは折り畳まれている場合です。その場合は”フィールドウェル”と書かれている部分をクリックして広げてください。

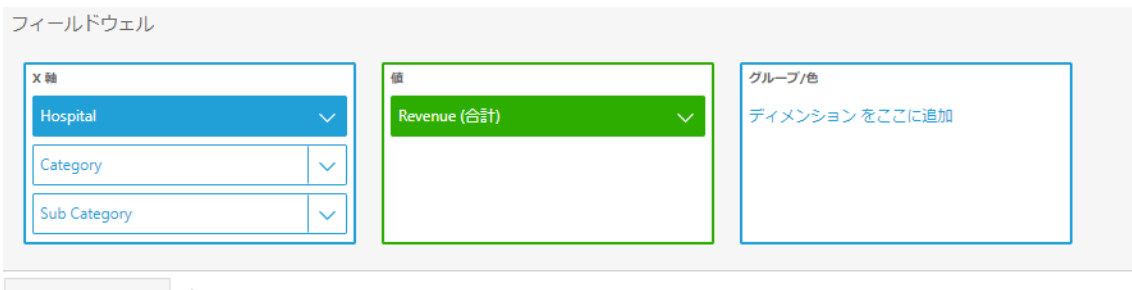
今回フィールドウェルには、”X 軸”、“値”、“グループ/色”が用意されます。ここでは、病院 (Hospital) ごとの利益 (Revenue) を見るので、まず、フィールドリストの Hospital をクリックしてください。自動的にフィールドウェルの”X 軸”に Hospital が挿入されます。これはフィールドの型や定義から自動的に判断してフィールドが追加される仕組みです。続いて、Revenue をクリックすると、”値”に Revenue(合計)として挿入されます。

これで病院ごとの Revenue の合計が得られました。これにドリルダウン（階層）を付けていきます。今回は Hospital の下階層として、Category を置き、そのさらに下に Sub Category を配置します。

フィールドリストから Category ドラッグ&ドロップ (D&D) で、X 軸にすでにある Hospital の下に移動させます。このとき、下図のように”ドリルダウンレイヤーの追加”と表示される場所でドロップしてください。

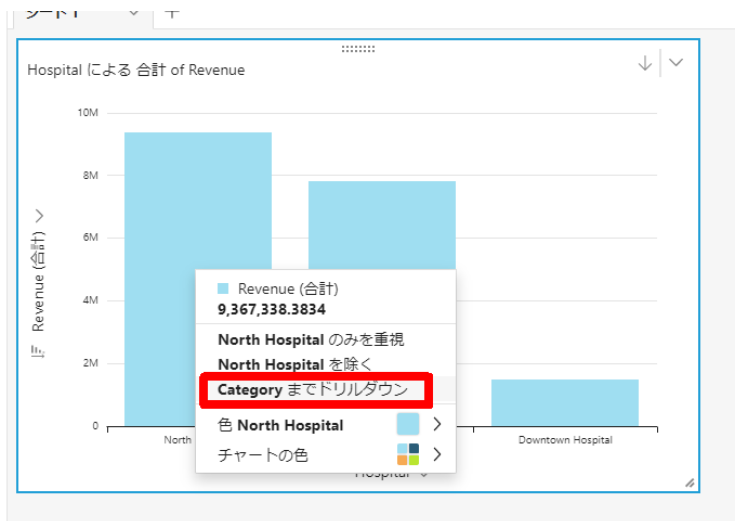


同様に、Category の下に、Sub Category をドリルダウンレイヤーとして D&D してください。
結果、フィールドウェルは以下ようになります。

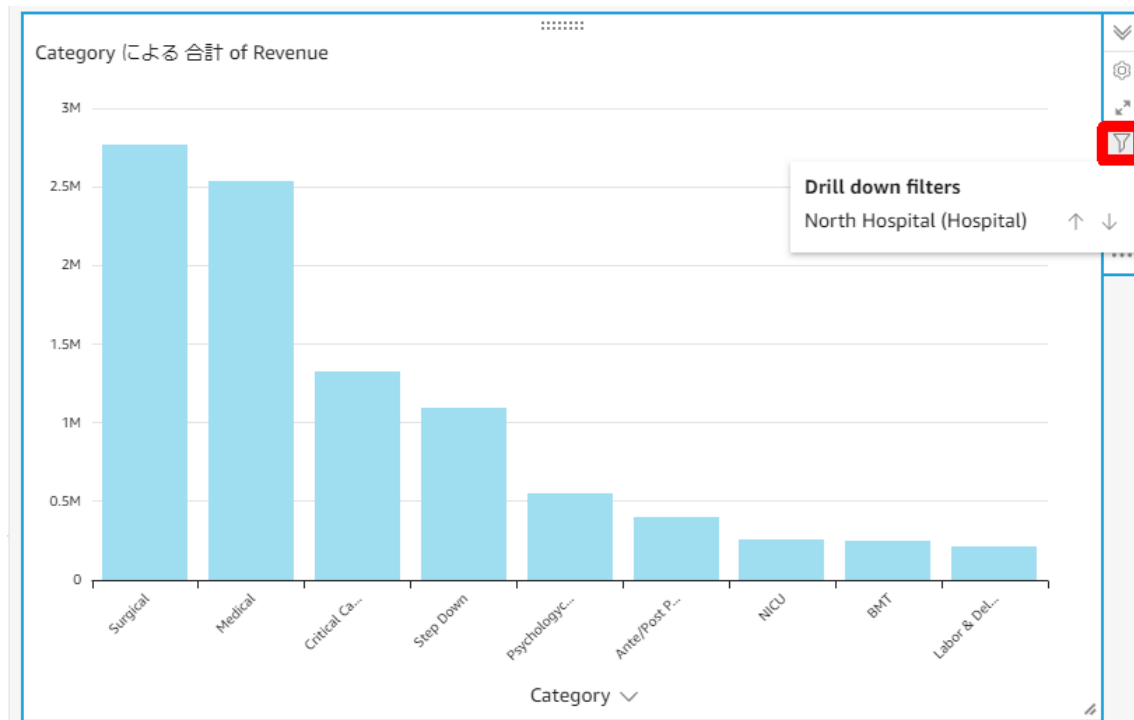


これでドリルダウンの階層が設定できましたので、実際に動かしてみましょう。

グラフの North Hospital をクリックし、メニューの”Category までドリルダウン”をクリックします。

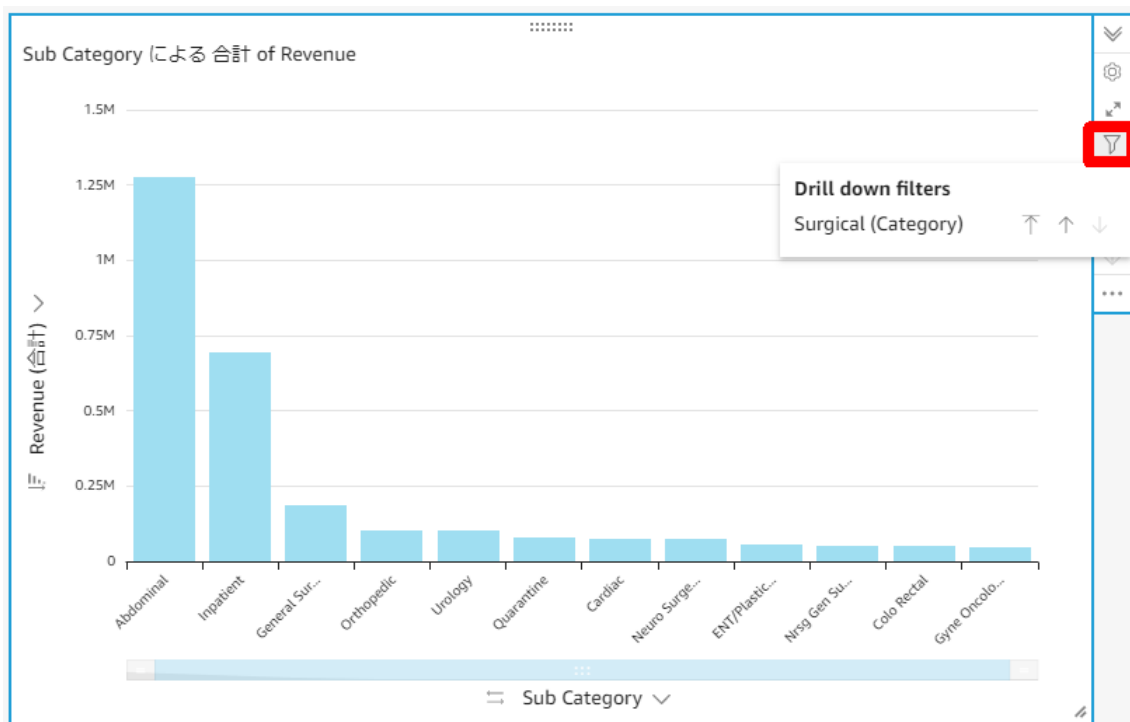


すると以下のように、“North Hospital”の中での Category 別の Revenue の表示に替わります。

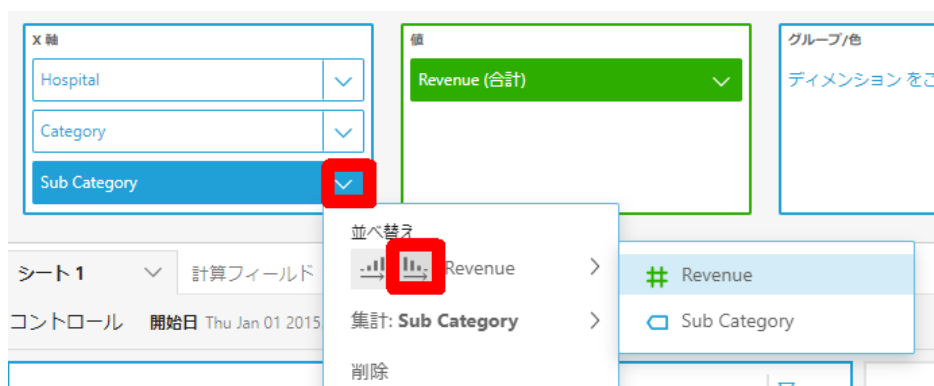


ビジュアル右側の斗マークをクリックすると、“North Hospital(Hospital)”と表示されます。これが North Hospital だけの範囲でドリルダウンが行われていることを示しています。

さらに一番左に表示されている“Surgical”をクリックし、“Sub Category までドリルダウン”をクリックします。これで、North Hospital において、Surgical 内での Sub Category ごとの Revenue が表示されます。



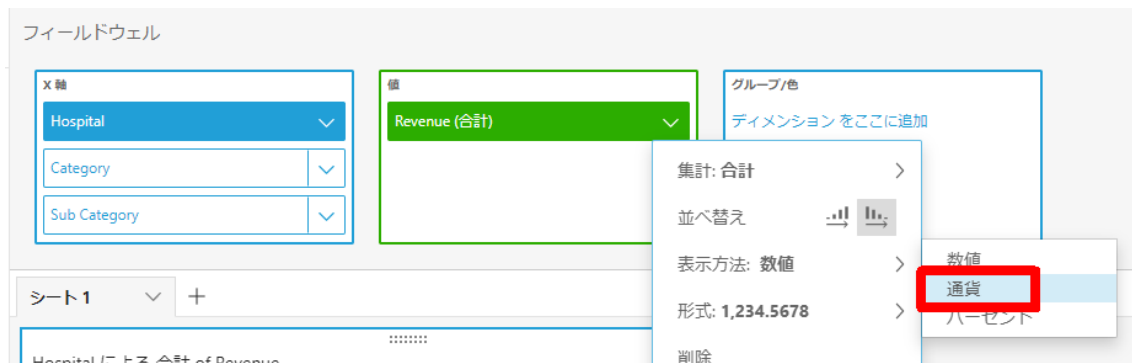
※補足上記の順でデータが並ばなかった場合は、下図のように Sub Category の並べ替えを大きい順に変更してください



ドリルアップをするには、ビジュアル右にある上向きの矢印（↑）をクリックします。また、↑を押すと最上位（Hospital）までドリルアップします。↑をクリックしてください。

ビジュアルの調整とコピー

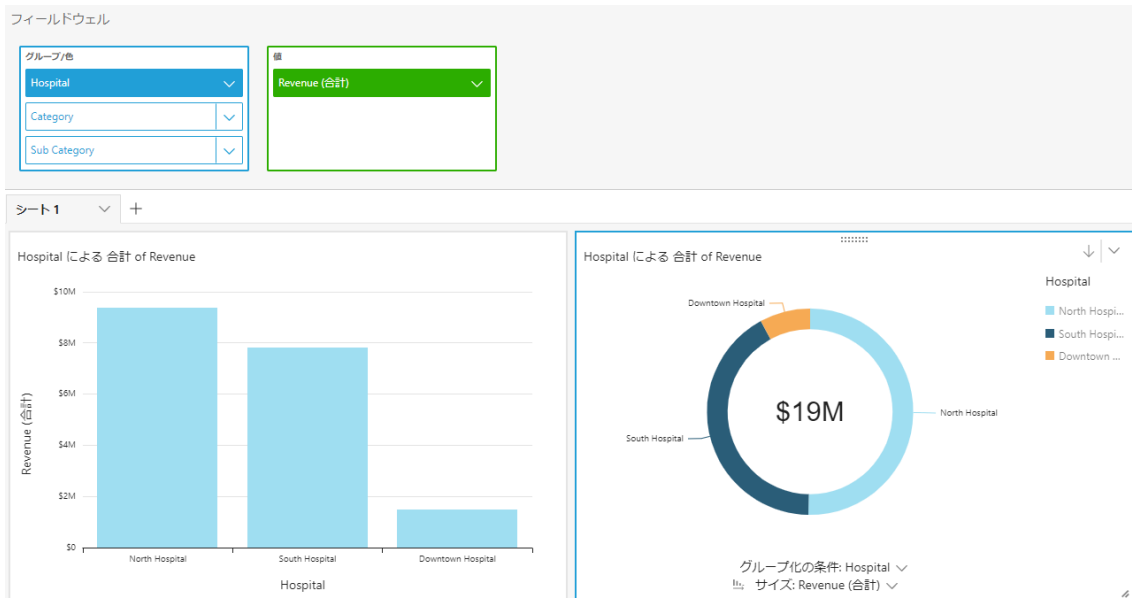
作成したビジュアルの”Revenue(合計)”は US ドルなのですが、画面上は単なる数値になっているため、フィールドウェルの”値”にある”Revenue(合計)”の右にある”V”をクリックし、表示方式を”通貨”にします。



これで画面上の数値に\$マークが表示されるようになります。

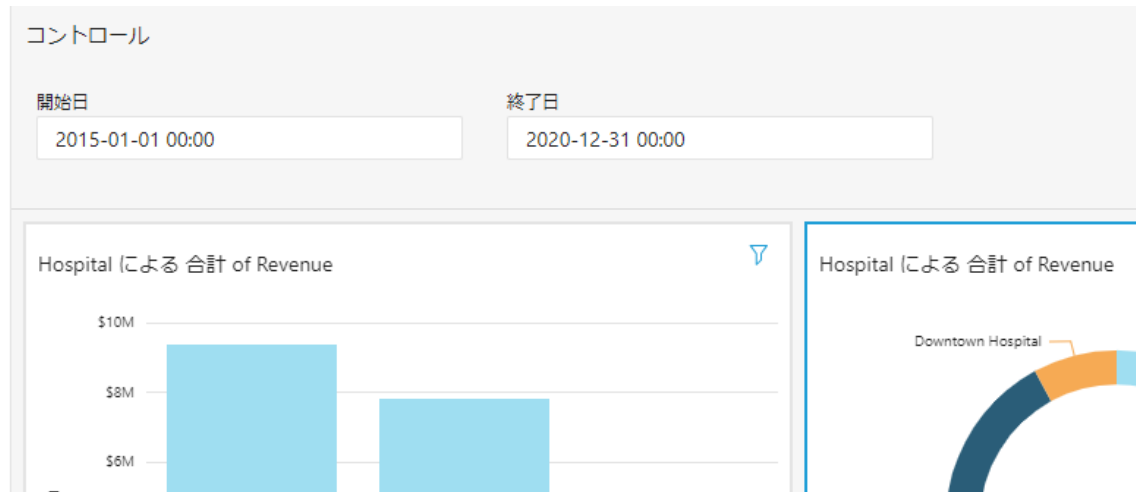
次にこの垂直棒グラフをコピーして別のビジュアルに変更します。ビジュアルの右側にある”…”をクリックして、”ビジュアルの複製”をクリックすると複製されます。表示の設定（先程の\$マーク）等もそのまま引き継がれていることが分かります。

この新しいビジュアルが選択された状態（新しく作られたビジュアルが青色線で囲まれている状態）で、左下のビジュアルタイプ一覧からドーナツグラフ(最上段、左から4つ目)を選択してください。以下のようにフィールドウェルの設定はそのままにビジュアルが変更されます。



コントロールの作成とフィルタ

ここでは、ここまでに作成したビジュアルに対して、ユーザが操作できるコントロールという GUI パーツ（下図）を画面上に作成し、データをフィルタ（絞り込み）可能にします。



コントロールは、Reader ユーザ（分析を作成する権限を持たない、読み取りのみのユーザ）でも操作できるため、コントロールを付けた分析を公開することで、インタラクティブなダッシュボードが実現できます。以下のような流れで作成します。

- ① 最初に**パラメータ**を作成します。これは変数のようなもので、ユーザが指定した値を保存する領域です
- ② 次に、①で作ったパラメータにリンクさせた、**コントロール**（GUI）を作成します。これにより GUI で入力した値がパラメータ（変数）に格納されます
- ③ 最後に、②で作成したパラメータを条件に指定した**フィルタ**を作成します。これにより GUI の操作がフィルタ条件に反映され、絞り込みが行われます

まずパラメータから作成します。画面左端の“パラメータ”をクリックし、“作成...”を選択します。パラメータには英数字の名前を付ける必要があります。今回は BeginDate としました。また、データタイプは“日時”にし、静的デフォルト値には 2015-01-01 を指定し、“作成”を押します。

新しいパラメータを作成

パラメータを使用して、フィールド、フィルター、シートの値を動的に制御します

名前

BeginDate

データタイプ (作成後は変更不可)

日時

静的デフォルト値

2015-01-01 00:00

動的デフォルト値を設定

キャンセル

作成

作成すると、続いてパラメータの接続先を選択するダイアログが出るので、コントロールを選択します。

パラメータが追加されました

パラメータを接続

パラメータ、新しいコントロール、フィルタの組み合わせを使用してフィルタを作成します。

フィルタ

フィルタまたは計算済みフィールドに向けた新しいコントロールを作成します。

コントロール

計算済みフィールドでパラメータを使用します。

計算フィールド

パラメータで URL アクションを作成します

カスタムアクション

閉じる

以下のダイアログでは、ユーザに見せるコントロールの名前を付けます。ここでは日本語を含む文字列を使用できます。今回は”開始日”としました。スタイルは規定の”日付選択ツール”のままにし、”追加”を押します。

パラメータにコントロールを追加

パラメータ
BeginDate

表示名

開始日

スタイル
日付選択ツール

キャンセル

追加

これで、ビジュアルの上方に”開始日”のコントロールが作成されます。（コントロールと書かれている部分をクリックして開いてください）

シート 1

コントロール

開始日
2015-01-01 00:00

Hospital による 合計 of Revenue
\$10M
\$8M

これで、1つ目のパラメータとコントロールは作成できました。日付を範囲でフィルタ（絞り込み）するには、もう一つ、終了日のパラメータとコントロールが必要です。左ペイン「パラメータ」の右側にある”+”を押し、以下の条件でもう一つ作成してください。

- パラメータ名：EndDate
- パラメータの型：日時型
- パラメータの静的デフォルト：2019-01-01
- コントロールの名前：終了日

以下ようになります。

パラメータ	+	フィールドウェル	X 軸	Hospital	値	Revenue (合計)	グループ/色
BeginDate	▼	シート1	▼	+			
EndDate	▼	コントロール					
		開始日			終了日		
		2015-01-01 00:00			2019-01-01 00:00		

これで2つのコントロール（開始日と終了日）が作成され、入力された値はそれぞれに紐付いたパラメータ（BeginDate と EndDate）に入るようになりました。次にパラメータを条件にしたフィルタを作成します。画面左端から”フィルタ”を選択、”作成...”を選択して、フィールドは”Admit Date”を選択します。

フィルタの編集 <

すべてのビジュアル ▼

Admit Date 
次の間 - なし

フィルタタイプ

期間 ▼

次の間 ▼

☒ パラメータを使用

開始日のパラメータ

BeginDate ▼

☒ 開始日を含める

終了日のパラメータ

EndDate ▼

☐ 終了日を含める

または

左側ペインに”Admit Date”のフィルタが作成されるのでクリックして開き、図のように設定します。

一番上にある選択で”すべてのビジュアル”を指定。これでフィルタが Analysis 全体のビジュアルに適用されます

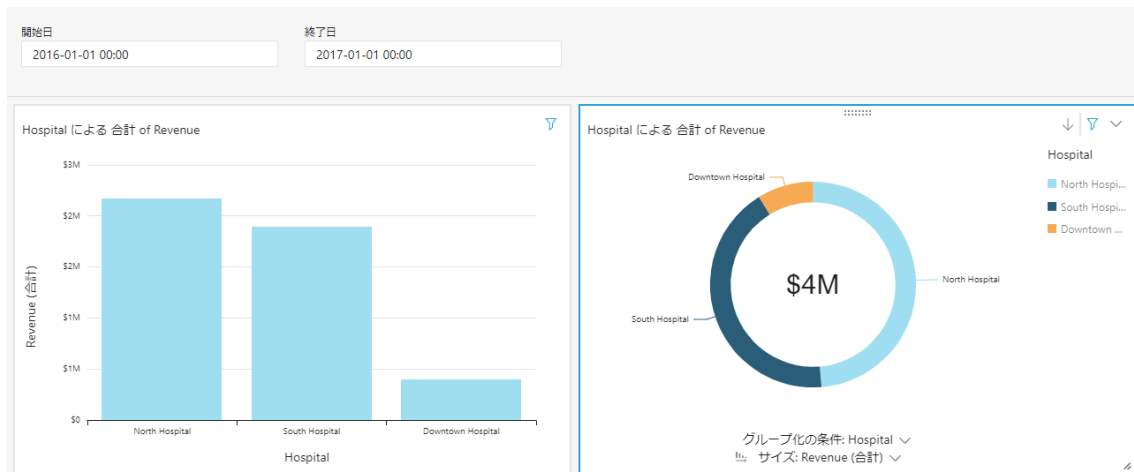
フィルタタイプは、期間、次の間を選択し、パラメータを使用にチェックを入れます。

開始日のパラメータには BeginDate を指定し、”開始日を含める”にチェックを入れます。チェックを入ると、指定した値を含んだ範囲指定になります。（＜と≦の違い）

終了日のパラメータには EndDate を指定し、”終了日を含める”のチェックは入れません。

設定できたら、”適用”をクリックします。

これで、コントロール（GUI）で入れた値がパラメータ（変数）に入り、それによりフィルタが実施される準備が整いました。開始日や終了日を変更し、2つのビジュアルがどう変わるかを確認してください。（例えば、開始日を 2016-01-01 に、終了日を 2017-01-01 に指定すると、2016 年のみのデータに絞り込まれます）



計算フィールド①：シンプルな計算フィールド

ここからは、計算フィールドの機能を使用した分析を作成します。計算フィールド(Calculated Field)とは、擬似的な列の機能です。データソースには無かった列を追加することで、より高度な分析が可能になります。

計算フィールドは、作成中の分析(Analyze)上に追加する方法と、データセット自体に追加する方法の2種類があります。本ハンズオンでは前者を使用します。後者のデータセット自体への追加の場合、利用できる関数等に制限がありますが、計算可能な部分は事前に計算しておき、表示時に毎回計算する必要を無くせるというメリットがあります。

まずはシンプルな計算フィールドを作成します。データを四半期（クォーター、Quarter）単位で分析したいと思っても、データセットにはクォーターフィールド（列）がありません。そこで計算フィールドで擬似的にクォーターフィールドを作成します。

日付自体は Admit Date に存在するため、その日付に extract 関数を使ってクォーターを抽出します。extract は日付から任意のフォーマットで文字列を抽出する関数で、抽出フォーマットを第1引数に、日付フィールドを第2引数に取ります。以下のように使用するとクォーター番号が抽出できます。

```
extract("Q",{Admit Date})
```

これを計算フィールドとして定義します。画面左上の”+追加”から、”計算フィールドを追加”を選択し、計算フィールド名に **Quarter**、計算式に上記式を入力し、”作成”を押します。

計算フィールドを編集する

関数リスト

- abs
- addDateTime
- avg
- avgIf
- avgOver
- ceil
- coalesce
- concat
- count

フィールドリスト

- Account Rep
- Admit Date
- Category
- Cost
- Discharge Date
- Discount
- Function
- Hospital

計算フィールド名

Quarter

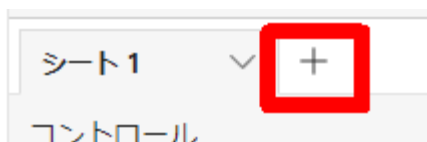
計算式

extract("Q",{Admit Date})

作成

これで Quarter という名前で計算フィールドが作成されました。フィールドリストにも Quarter というフィールドが表示されているはずです。この Quarter フィールドを使ってピボットテーブルを作成します。

新しいピボットテーブルを作成する前に、シートを切り替えます。シート名（シート 1 等）の横にある“+”を押すと、新しいシートが作成されます。



シートにはそれぞれ任意の名前を付けることができます。新しいシートの中に 1 つビジュアルが作成されていますので、ビジュアルタイプからピボットテーブル（上から 4 段目、左から 2 つ目）を選択します。ピボットテーブルのフィールドウェルには、以下のようなよう、クリックや D&D でフィールドを設定してください。

フィールドウェル

行

- Admit Date
- Quarter

列

ディメンションをここに追加

値

Revenue (合計)

次にフィールドウェル上で、Admit Date の右側にある”V”をクリックして表示を調整します。集計を”年”に設定し、さらに形式は”2020”のように年だけの表示にしてください。同様に”値”にある”Revenue (合計)”は、表示方法を”通貨”に変更します。

以下のようになります。

フィールドウェル

行	列	値
Admit Date (YEAR) ▼	ディメンションをここに追加	Revenue (合計) ▼
Quarter ▼		

シート 1 計算フィールド ▼ +

Admit Date and Quarter による 合計 of Revenue

Admit Date	Quarter	Revenue
2015	1	\$1,561,877.72
	2	\$1,092,500.97
	3	\$1,271,644.94
	4	\$1,311,380.01
2016	1	\$1,032,065.62
	2	\$1,060,424.57
	3	\$1,076,119.85
	4	\$1,282,178.57

年単位で Admit Date が表示され、その横に Quarter フィールドが第 1 四半期から第 4 四半期まで表示されている事が分かります。値として四半期毎の Revenue(合計)が表示されています。

計算フィールド②：テーブル計算関数（RUNNINGSUM）

作成したピボットテーブルでは四半期毎の Revenue がありますが、これに加えて、各四半期の**その年の“累計”**を表示するようにします。つまり、第 2 四半期は第 1 四半期 + 第 2 四半期の値を、第 3 四半期には、第 1 四半期 + 第 2 四半期 + 第 3 四半期の値という累計を新たな値として追加します。

こういった“ある期間ごとの集計”の機能は、QuickSight では**テーブル計算関数**として提供されており、累計には runningSum 関数が用意されています。（※関数一覧等、ドキュメントへのリンクは資料末尾にあります）

runningSum は第 1 引数に、集計する値を（この場合は Sum(Revenue)）、第 2 引数にはソート順の指定、第 3 引数にはパーティション（今回は年単位）を指定します。

runningSum(sum(Revenue),[{Admit Date} ASC],[{Admit Date}])

今回は上記の計算式で、**YearToDateSales** という名前の計算フィールドを作成してください。

（※フィールド名は大文字小文字を区別します。この後、この計算フィールドを別の計算フィールド内で使用するため、名前の打ち間違いに注意してください）

作成したら、値としてフィールドウェルに追加し、表示方法を“通貨”にします。ピボットテーブルは以下のようになります。

フィールドウェル

行

Admit Date (YEAR) ▼

Quarter ▼

列

ディメンションをここに追加

値

Revenue (合計) ▼

YearToDateSales (カスタム) ▼

シート 1 計算フィールド ▼ +

Admit Date and Quarter による 合計 of Revenue and Yeartodatesales

Admit Date	Quarter	Revenue	YearToDateSales
2015	1	\$1,561,877.72	\$1,561,877.72
	2	\$1,092,500.97	\$2,654,378.69
	3	\$1,271,644.94	\$3,926,023.63
	4	\$1,311,380.01	\$5,237,403.64
2016	1	\$1,032,065.62	\$1,032,065.62
	2	\$1,060,424.57	\$2,092,490.20
	3	\$1,076,119.85	\$3,168,610.04
	4	\$1,282,178.57	\$4,450,788.61
2017	1	\$1,058,493.45	\$1,058,493.45
	2	\$935,728.42	\$1,994,221.88
	3	\$1,048,166.83	\$3,042,388.70
	4	\$1,252,884.08	\$4,295,272.78
2018	1	\$1,200,943.33	\$1,200,943.33

YearToDate 列がその四半期までの Revenue の積算になっているのが確認できます。

(オプション) 計算フィールド③: テーブル計算関数
(PERCENTDIFFERENCE)

ピボットテーブルに、昨年同四半期との比較を追加します。例えば 2018 年第 3 四半期時点の累積レベニューは、2017 年第 3 四半期時点と比較して何%の違いがあるかといった計算です。こういった計算にはテーブル計算関数の percentDifference を使用します。

percentDifference は第 1 引数に、比較する値、第 2 引数にソート基準、第 3 引数にインデックスを取ります。インデックスとは、相対的にどの位置にある値を比較対象にするかを指します。以下の式を見てください。

percentDifference(YearToDateSales, [{Admit Date} ASC], -4, [])

インデックスに -4 と書かれています。つまり **4 つ前**のデータを比較対象にします。今回のピボットテーブルは四半期単位で行を構成しているため、4 つ前は 1 年前という事になります。

※補足: このように各テーブル計算関数はデフォルトでは、**表示された時点のデータの並びに依存した形**で動作しているという点に注意してください。つまり、-4 が 1 年前のデータを指すのは、あくまで今回のピボットテーブルが四半期単位で集約されているためです。表示や集約の結果に依存せずに演算を行う機能としては LAA (Level-Aware Aggregation) が用意されています。LAA はオプションのハンズオンで解説しています。

ではこれまでと同様に、上記の式を **YTD % Difference** という名前の計算フィールドとして定義し、フィールドウエルの”値”の一番下に追加します。今回は比率なので、表示方法は”パーセント”を選択します。

以下のような表示になります。

行

Admit Date (YEAR) ▼

Quarter ▼

列

ディメンション をここに追加

値

Revenue (合計) ▼

YearToDateSales (カスタム) ▼

YTD % Difference (カスタム) ▼

行 列

シート 1

計算フィールド ▼

+

.....

Admit Date and Quarter による 合計 of Revenue, Yeartodatesales, and Ytd % Difference

Admit Date	Quarter	Revenue	YearToDateSales	YTD % Difference
☐ 2015	1	\$1,561,877.72	\$1,561,877.72	
	2	\$1,092,500.97	\$2,654,378.69	
	3	\$1,271,644.94	\$3,926,023.63	
	4	\$1,311,380.01	\$5,237,403.64	
☐ 2016	1	\$1,032,065.62	\$1,032,065.62	-33.92%
	2	\$1,060,424.57	\$2,092,490.20	-21.16%
	3	\$1,076,119.85	\$3,168,610.04	-19.29%
	4	\$1,282,178.57	\$4,450,788.61	-15.01%
☐ 2017	1	\$1,058,493.45	\$1,058,493.45	2.56%
	2	\$935,728.42	\$1,994,221.88	-4.69%
	3	\$1,048,166.83	\$3,042,388.70	-3.98%
	4	\$1,252,884.08	\$4,295,272.78	-3.49%
☐ 2018	1	\$1,200,943.33	\$1,200,943.33	13.45%
	2	\$1,166,292.13	\$2,367,235.46	18.7%
	3	\$1,073,458.86	\$3,440,694.32	13.09%
	4	\$1,209,204.94	\$4,649,899.27	8.25%

2015 年は、その前の年がデータに存在しないため表示されていませんが、2016 年からは前年同四半期時点での累積としての差が計算されています。

(オプション) 計算フィールド④: テーブル計算関数 (PERCENTOFTOTAL)

ピボットテーブルに、各四半期のレベニューが、その年の売上の何%を占めているかの情報を追加します。この場合、関数 `percentOfTotal` が利用できます。第 1 引数に演算式、第 2 引数にパーティショニングの基準（この例では、1 年ごと）を指定します。

```
percentOfTotal(sum(Revenue),[{Admit Date}])
```

これまでと同様に上記の式を、**PercentOfTotalSales** という計算フィールド名で定義し、“値”のフィールドウェル一番下に追加し、表示方法を“パーセント”に変更してください。以下のようになります。PercentOfTotalSales 列で、各四半期がその年のレベニューの何%を占めているかが表示できました。

行

Admit Date (YEAR) ▼

Quarter ▼

列

ディメンション をここに追加

値

Revenue (合計) ▼

YearToDateSales (カスタム) ▼

YTD % Difference (カスタム) ▼

PercentOFTotalSales (カスタム) ▼

シート 1 ▼

+

Admit Date and Quarter による 合計 of Revenue, Yeartodatesales, Ytd % Difference, and Percentoftotalsales

Admit Date	Quarter	Revenue	YearToDate...	YTD % Diffe...	PercentOFT...
2015	1	\$1,561,877.72	\$1,561,877.72		29.82%
	2	\$1,092,500.97	\$2,654,378.69		20.85%
	3	\$1,271,644.94	\$3,926,023.63		24.28%
	4	\$1,311,380.01	\$5,237,403.64		25.03%
2016	1	\$1,032,065.62	\$1,032,065.62	-33.92%	23.18%
	2	\$1,060,424.57	\$2,092,490.20	-21.16%	23.82%
	3	\$1,076,119.85	\$3,168,610.04	-19.29%	24.17%
	4	\$1,282,178.57	\$4,450,788.61	-15.01%	28.8%
2017	1	\$1,058,493.45	\$1,058,493.45	2.56%	24.64%
	2	\$935,728.42	\$1,994,221.88	-4.69%	21.78%
	3	\$1,048,166.83	\$3,042,388.70	-3.98%	24.4%
	4	\$1,252,884.08	\$4,295,272.78	-3.49%	29.16%
2018	1	\$1,200,943.33	\$1,200,943.33	13.45%	25.82%
	2	\$1,166,292.13	\$2,367,235.46	18.7%	25.08%
	3	\$1,073,458.86	\$3,440,694.32	13.09%	23.08%
	4	\$1,209,204.94	\$4,649,899.27	8.25%	26%

ML インサイトによる異常値検出の設定

ここからは、QuickSight の ML(機械学習)インサイトの機能のうち、異常値検出とナラティブを体験します。

まず ML インサイトのハンズオン用に、シートの横にある”+”ボタンを押し、別のシートを作成します。シートには任意の名前を付けることができます。



新しいシートには、空白のビジュアルが 1 つ用意されていますが、ここでは使わないため、ビジュアルの右の”…”を押し、”削除”を選択してください。

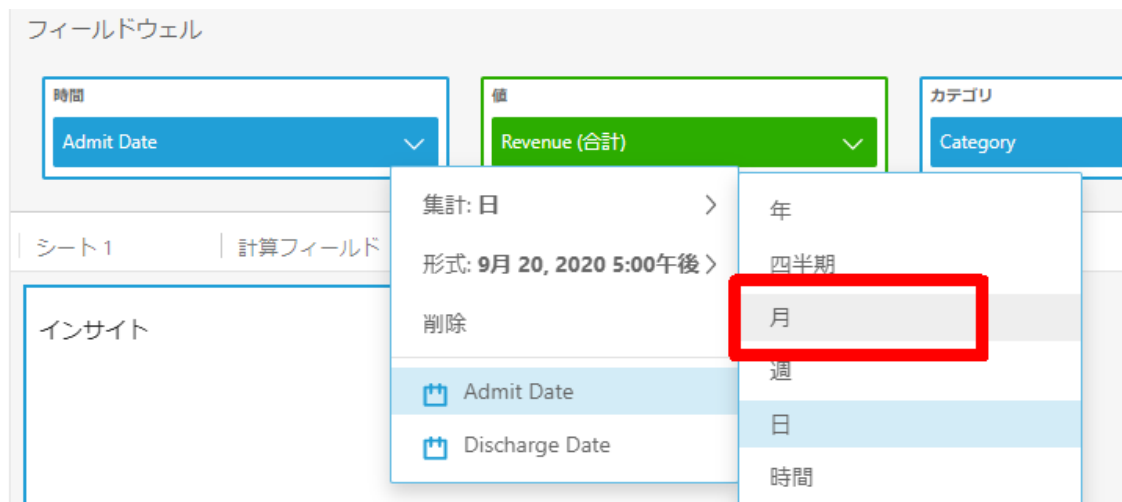
まず、異常検知の計算を設定します。この**異常値検出の機能を利用するにはユーザ料金とは別に料金が必要**になります。 (<https://aws.amazon.com/jp/quicksight/pricing/>)

ただし、QuickSight をサインアップして最初の 2 ヶ月間は、1 ヶ月あたり最大 500,000 メトリクス分の無料枠があるため、今回初めてサインアップした方は無料枠に十分収まります。

画面左上の”+追加”から、”インサイトを追加”を選択し、”異常値検出”を選択し、”選択”を押します。”時間”、”値”、”カテゴリ”のフィールドウェルが用意されますので、それぞれ、以下のようにフィールドを配置してください。



この際、“時間”の Admit Date は以下のように”集計:月”を必ず選択してください。（月単位にしないと、計算量が増えてしまうため）



次にインサイトの”今すぐ始める”を押します。

設定はほとんどデフォルトのままで問題ありませんが、“寄与率の分析”は設定を変更します。これは、異常値を検出した際、その異常がどのフィールドの組み合わせによって起こったのかという計算（寄与率の算出）を行うものです。対象とするフィールドは最大4つまで指定できるので、ここでは以下のように、Patient ID、Payer Segment、Physician、Service の4つを選択します。

異常検出の設定

スケジュール開始日:
2020-03-05 22:09

タイムゾーン
Asia/Tokyo

寄与率の分析 (オプション)
異常の原因を理解することができます。

Patient ID, Payor Segment, Physician, Service

Hospital	<input type="checkbox"/>
Patient ID	<input checked="" type="checkbox"/>
Payor Segment	<input checked="" type="checkbox"/>
Physician	<input checked="" type="checkbox"/>
Priority	<input type="checkbox"/>
Region	<input type="checkbox"/>
Service	<input checked="" type="checkbox"/>
State	<input type="checkbox"/>
Sub Category	<input type="checkbox"/>
Zip Code	<input type="checkbox"/>

絶対値 パーcentage

削除 キャンセル 保存

上記の状態ですべての項目にチェックを入れ、インサイトの画面にある「今すぐ実行」を押すと、異常値を検出するための計算が始まります。今回の設定の場合、演算には約 2-3 分間かかります。これを待つ間に、もう一つのナラティブ機能を確認します。

ML インサイトによるナラティブ

ナラティブは、インサイトを図ではなく人間の言葉や箇条書きで表現する機能です。

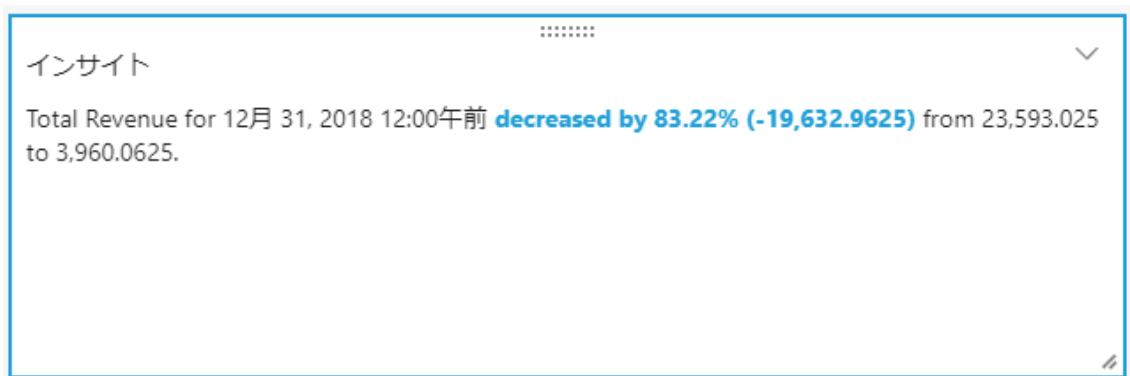
画面左上の”+”から”インサイトを追加”を選択し、”前期比”を選択し、”選択”を押します。これは、一定期間前と比較した現在の状況を説明するナラティブ（文章）です。

フィールドウェルが表示されるので、以下のように設定してください。

フィールドウェル

時間	値	カテゴリ
Admit Date	Revenue (合計)	ディメンションをここに追加

すると、以下のような結果が得られます。



このサンプルデータは最新の日時が 2018 年 12 月 31 日のため、31 日と、その前日と比較した数字が文章で表示されているのが分かります。これがナラティブ機能です。

ナラティブはデフォルトで多数のプリセットが用意されていますが、自分でも文章をカスタマイズする事が可能です。インサイト右の”...”をクリック⇒”説明をカスタマイズ”を選択してください。ナラティブエディタが表示されます。

説明を編集

キャンセル保存

コードを挿入

12パラグラフ

B

i

U

↶

Abc

Abc

:

:

:

:

☐

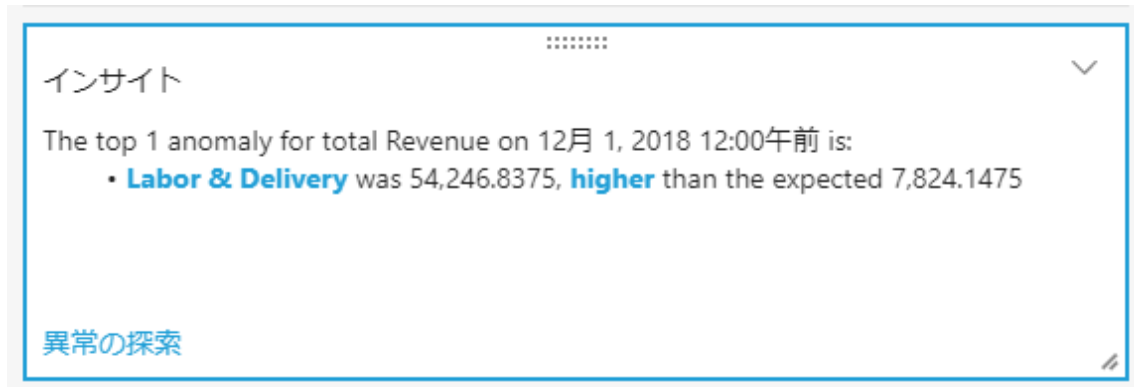
Total PeriodOverPeriod.metricField.name for PeriodOverPeriod.currentTimeValue.formattedValue If PeriodOverPeriod.percentDifference.value >= 0 increased by If PeriodOverPeriod.percentDifference.value < 0 decreased by PeriodOverPeriod.percentDifference.formattedAbsoluteValue (PeriodOverPeriod.absoluteDifference.formattedValue) from PeriodOverPeriod.previousMetricValue.formattedValue to PeriodOverPeriod.currentMetricValue.formattedValue

GUI のリッチテキストエディタで、中に式や IF 文が入っているのが確認できます。ナラティブの文章はこのような形でロジックを組むことで実現されています。

今回は編集せず、右上の”キャンセル”を押して、前の画面に戻ります。

ML インサイトによる異常値検出の内容を確認する

先に実行しておいた異常値検出の計算が完了すると以下のような画面になっているはずです。



インサイトでは1つの異常（Anomaly）が12月に発見されたとあり、内容としては Labor & Delivery において、想定（予想）の 7,824 よりずっと高い（異常に高い）54,236 を記録したと書いてあります。このように、フィールドの色々な組み合わせから、異常と思われる値を探し出すのが異常値検出の目的です。

インサイトの左下にある”異常の探索”を押して、詳細画面に移動します。



右のグラフは異常値を表しています。計算時に”月単位”にしたため、異常値の計算は月あたり一回ずつになっています。Labor & Delivery は 12 月に急に増加しているのがこのグラフからも分かります。

また、左側ペインには、寄与率が掲載されています。この異常な増加の要因となったのが、Patient ID 40032 と、7427 であること等、影響が大きかったものが自動的に記載されています。

内容を確認したら、上部の”分析に戻る”で前の画面に戻ります。

これでハンズオン 2 は完了です。

(オプション) ハンズオン 3 : LEVEL-AWARE AGGREGATION

LEVEL-AWARE AGGREGATION とは

本章では、Level-Aware Aggregation(LAA)機能を使った高度な分析を体験します。LAA は関数の計算（集計等）を行う**タイミング**を指定する機能です。

テーブル計算関数のオプションハンズオンで、関数が「ビジュアルが表示されたタイミング」に依存して動いていることが確認しましたが、このような関数が実行されるタイミングを調整する機能が LAA です。（LAA が利用できる関数一覧については、ドキュメントを確認してください。資料末尾にドキュメントへのリンクがあります）

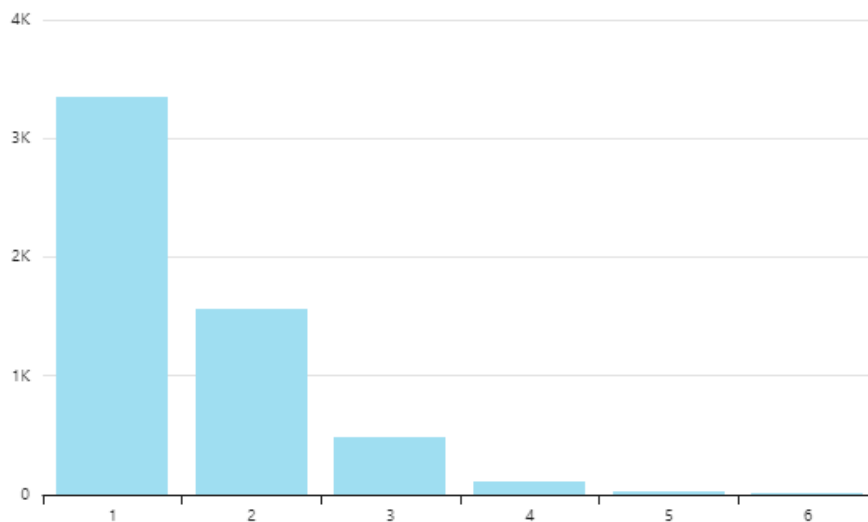
以下のタイミングを指定可能です。

- PRE_FILTER – データセットをフィルタする前に計算
- PRE_AGG – 集計をビジュアルに適用する前に計算
- POST_AGG_FILTER – ビジュアルが表示されるタイミングで計算（ここまで見てきた挙動であり、関数の**デフォルト動作**）

これがどのような場合に役に立つのかをハンズオンで確認します。LAA 用のハンズオンを開始する前に、シートを切り替えてください。シート名の横にある”+”を押すと、新しいシートが作成されます。

LEVEL-AWARE AGGREGATION (COUNTOVER)

以下の図のように、1 回入院した患者の総数は XX 名、2 回入院した患者の総数は XX 名というように、入院回数ごとの患者数をカウントしたいとします。この場合 X 軸にあたる「入院回数」はデータに無いため、入院患者数を集計する前に、X 軸を計算しておく必要があります。このようなケースでは PRE_AGG を指定して計算フィールドを作成します。



以下の内容で計算フィールドを作成してください

名前: **AdmittancePerPatient**

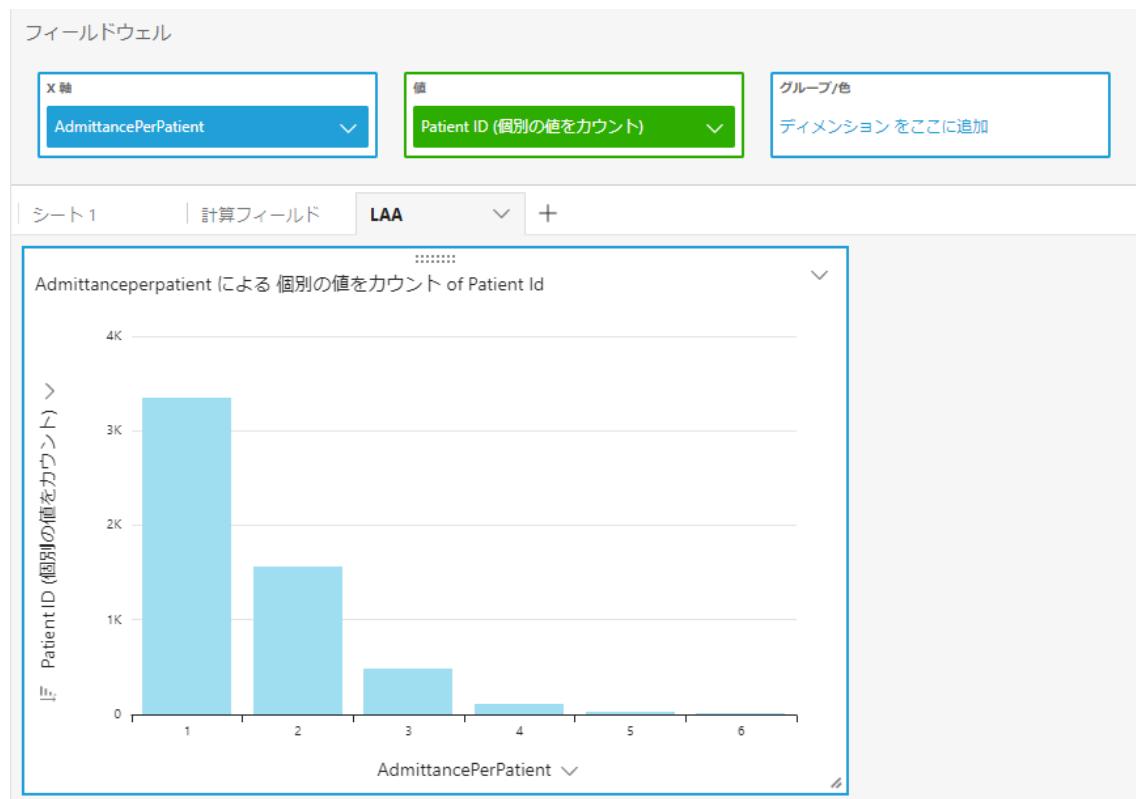
計算式: `countOver(({Admit Date}),[{Patient ID}],PRE_AGG)`

`countOver` の第 1 引数は、カウントの対象ですが、今回は入院日のフィールドにしています。第 2 引数がパーティショニングで、上記では Patient ID (患者 ID) を指定していますので、患者 ID ごとに入院回数がカウントされます。第 3 引数で、PRE_AGG を指定しています。

AdmittancePerPatient フィールドが作成できたら、垂直棒グラフ (上から 2 番目、左から 2 番目) ビジュアルを作成し、フィールドウェルの X 軸に AdmittancePerPatient を、値に、

Patient ID を置きます。Patient ID は集計で”個別の値をカウント”を選択してください（SQL でいうところの distinct 演算）。

以下のようになれば完成です。



LEVEL-AWARE AGGREGATION (SUMOVER)

以下の図のように、病院(Hospital)ごとのレベニュー (Revenue) を集計したとします。

Hospital	Revenue
Downtown Hospital	1,480,027.1607
North Hospital	9,367,338.3834
South Hospital	7,785,998.76

上記の表をフィルタして Revenue を構成する患者(Patient)の中で、生涯 (通算) で\$10,000 以上支払っている Patient だけに絞りたい場合はどうすれば良いのでしょうか。

上記の表を可視化する時点ではすでにデータは集約されてしまっていますので、データを集約する前(PRE_AGG)で、Patient ごとの通算 Revenue を計算する必要があります。

まずは上記の表を表示します。新しいビジュアルを作成し、ピボットテーブル (上から 4 つ目、左から 2 つ目) を選択します。行には”Hospital”を、値には”Revenue“を設定し、Revenue は通貨に指定します。

フィールドウェル

行 Hospital	列 ディメンションをここに追加	値 Revenue (合計)
---------------	--------------------	-------------------

Hospital による 合計 of Revenue

Hospital	Revenue
Downtown Hospital	\$1,480,027.16
North Hospital	\$9,367,338.38
South Hospital	\$7,785,998.76

次に、Patient 毎の総支払い表す計算フィールドを LAA で定義します。

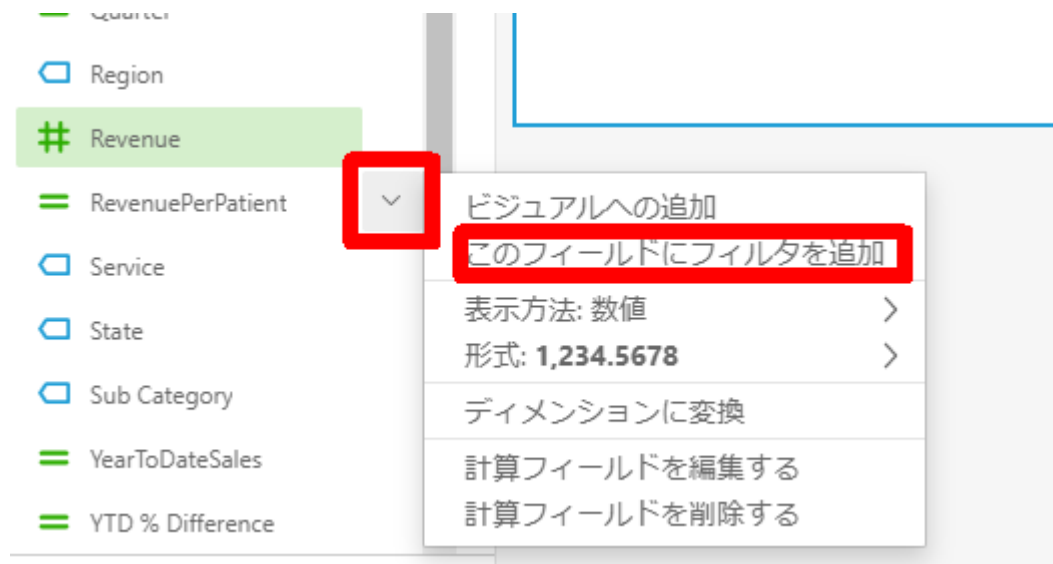
以下の内容で計算フィールドを作成してください。

名前： **RevenuePerPatient**

計算式： **sumOver(Revenue,[{Patient ID}],PRE_AGG)**

このようにすることで、Revenue が Patient ID 毎に合計されます。これをビジュアルのフィルタ条件に設定します。

先程作成したピボットテーブルが選択された状態で、フィールドリストの”RevenuePerPatient”の上にマウスカーソルを移動させ、横にある”V”をクリックし、“このフィールドにフィルタを追加”をクリックします。



作成されたフィルタで、以下のように 10,000 以上の値を残すように設定します。

フィルタの編集

このビジュアルのみ

RevenuePerPatient

次と等しい - なし

集計なし

次以上

☐ パラメータを使用

10000

Null を含む

これを適用すると、最終的な結果が得られます。

Hospital による 合計 of Revenue	
Hospital	Revenue
Downtown Hospital	\$439,359.30
North Hospital	\$3,600,224.28
South Hospital	\$4,838,198.60

まとめと後片付け

QuickSight の費用について

QuickSight 1 ユーザ分は無料で利用できますので、そのまま使っていただいても費用は発生しません。ただし、ハンズオン 2 において「異常値検出」を作成した場合、その異常値の演算には費用が発生します。例えばスケジュールを「毎日」に設定した場合は毎日演算が自動実行されます。異常値検出用に作成したインサイトの右にある“…”メニューから「削除」を選択することで削除できます。

QuickSight の利用全体を完全に削除する場合は QuickSight のトップ画面でリージョンをバージニア北部に変更した後、右上のユーザアイコン→QuickSight の管理→アカウントの設定を選択し、“サブスクリプション解除”で削除が可能です。（作成したデータセットや SPICE に投入したデータは全て削除されます）

補足情報・参考資料（QUICKSIGHT）

- ホームページ <https://aws.amazon.com/jp/quicksight/>
- ドキュメント <https://aws.amazon.com/jp/documentation/quicksight/>
- サービス全体の説明資料
 - <https://www.slideshare.net/AmazonWebServicesJapan/20180228-aws-black-belt-online-seminar-quicksight-89889593>
 - <https://www.slideshare.net/AmazonWebServicesJapan/20180801-aws-black-belt-online-seminar-amazon-quicksight>
 - <https://www.slideshare.net/AmazonWebServicesJapan/20200204-aws-black-belt-online-seminar-amazon-quicksight>
- 計算フィールド、表計算関数
 - <https://aws.amazon.com/jp/blogs/news/advanced-analytics-with-table-calculations-in-amazon-quicksight/>
 - https://docs.aws.amazon.com/ja_jp/quicksight/latest/user/working-with-calculated-fields.html
- Level-Aware Aggregation(LAA)関連
 - <https://aws.amazon.com/jp/blogs/news/create-advanced-insights-using-level-aware-aggregations-in-amazon-quicksight/>
 - <https://docs.aws.amazon.com/quicksight/latest/user/level-aware-aggregations.html>

■その他の AWS サービスについて

その他、AWS の各種サービスについての資料は以下にまとまっています。

<https://aws.amazon.com/jp/aws-jp-introduction/>