

データサイエンティスト講座 プレゼンテーション

郷田郁子 2020/11/29



今回の分析をするための方針

- 1) 季節性で予測数値が出やすいもの： アイスクリーム、ビール
- 2) 季節イベント性のあるもの： うなぎのかば焼き、もち
- 2) 変動してるけど理由がわからないもの： (変動してる→売上データの目視)
粉ミルク、乾うどん・そば、プリン、コロッケ

コロッケで全く傾向が見えない→ビールで実施



目的:ビールの需要を予測する

- ビールは季節により売上変動があると予測しました
- また、第3のビールやチューハイ人気などの飲酒トレンドの影響もあると想定しました。
- 需要予測を行い、適切な在庫・発注管理を行う効果が大いと考えました。

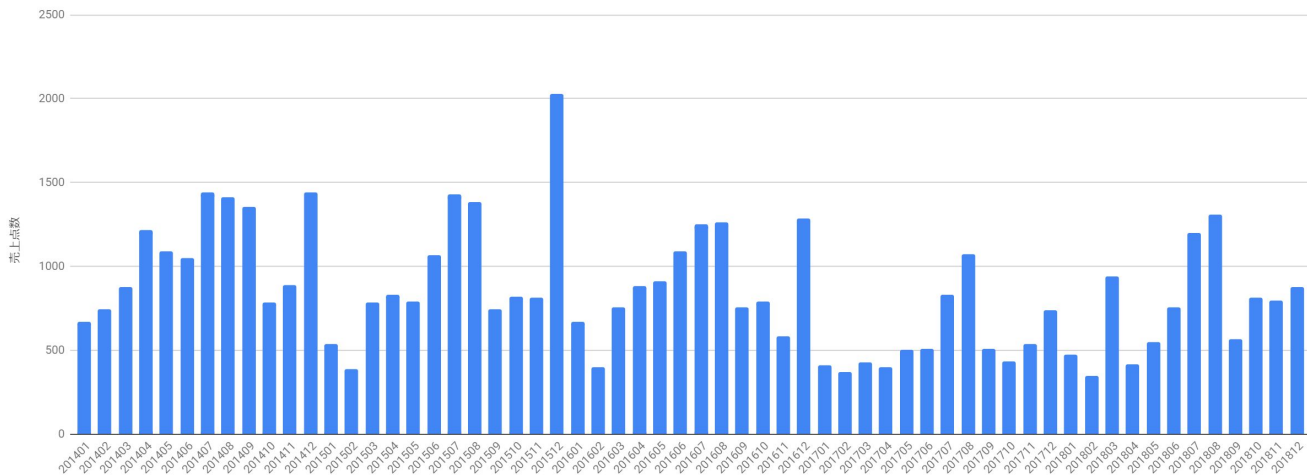


ビールの売上(月推移)

- ・定期的な変動と、全体的に右肩下がりの様子が見えます。

[グラフをコピー](#) [グラフを編集](#) [グラフを公開](#) [グラフをダウンロード](#) [グラフを削除](#) [^](#)

売上点数



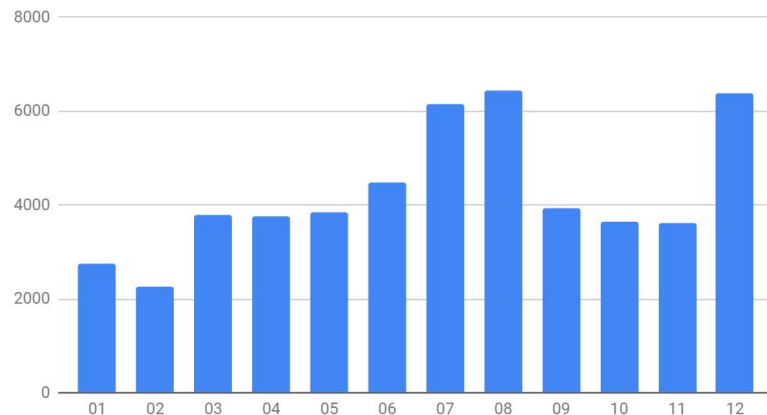
平均月売上数(2014~2018)

5年間の月の平均推移を見てみると

7月・8月・12月の売上数が多いです。

| 売上点数のSUM 月 | | | | | | | | | | | | | |
|------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|--------------|
| 年 | 01 | 02 | 03 | 04 | 05 | 06 | 07 | 08 | 09 | 10 | 11 | 12 | 総計 |
| 2014 | 668 | 746 | 876 | 1216 | 1089 | 1050 | 1441 | 1414 | 1355 | 784 | 891 | 1440 | 12970 |
| 2015 | 540 | 387 | 783 | 830 | 789 | 1066 | 1429 | 1382 | 743 | 819 | 816 | 2026 | 11610 |
| 2016 | 668 | 398 | 757 | 884 | 910 | 1090 | 1253 | 1261 | 757 | 792 | 585 | 1286 | 10641 |
| 2017 | 411 | 370 | 430 | 399 | 505 | 506 | 830 | 1072 | 506 | 434 | 540 | 741 | 6744 |
| 2018 | 472 | 349 | 942 | 417 | 551 | 755 | 1198 | 1310 | 567 | 813 | 795 | 878 | 9047 |
| 総計 | 2759 | 2250 | 3788 | 3746 | 3844 | 4467 | 6151 | 6439 | 3928 | 3642 | 3627 | 6371 | 51012 |

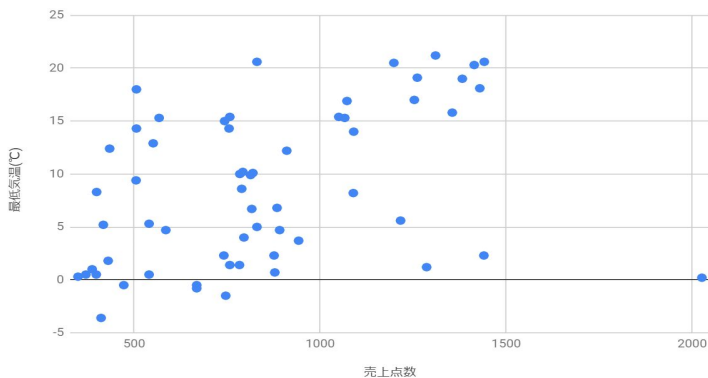
平均月売り上げ数



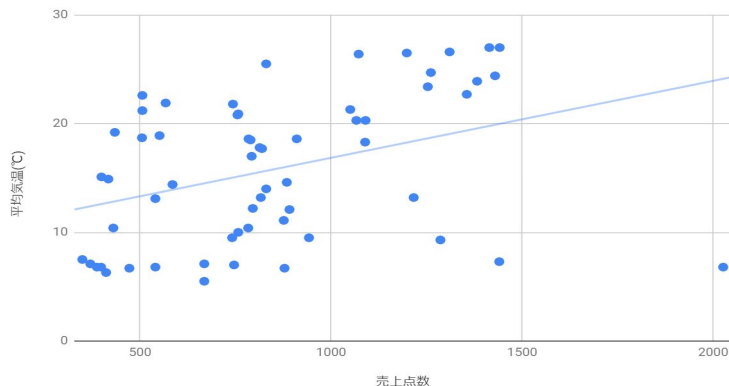
気温との関連性

気温との連動については大体の傾向はあるものの明確な連動があるようには見えませんでした。

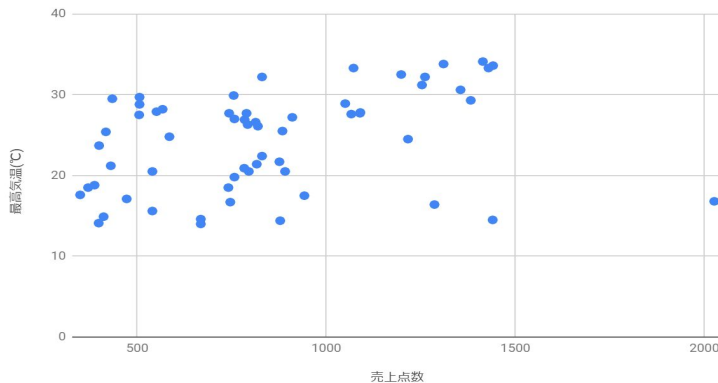
最低気温(°C) と 売上点数



平均気温(°C) と 売上点数



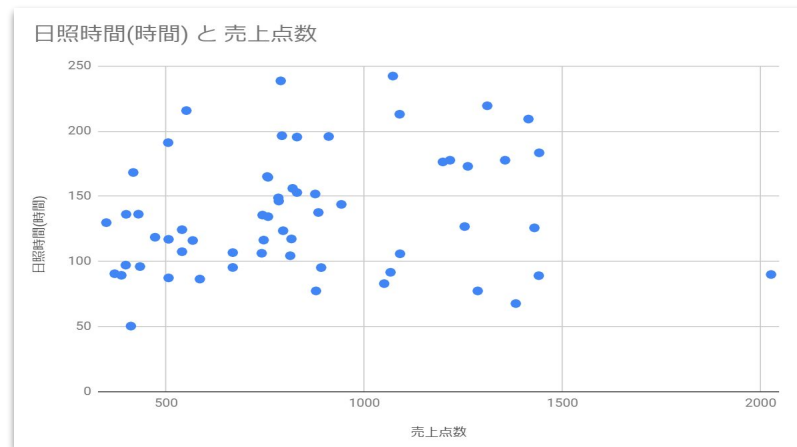
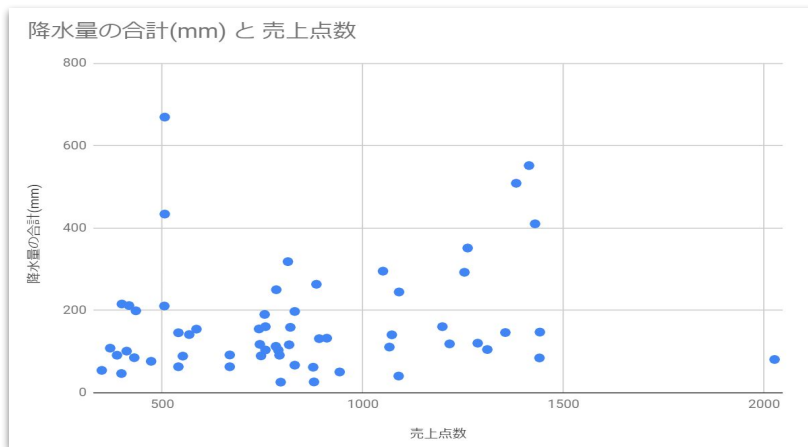
最高気温(°C) と 売上点数



天候と販売数の関連性

降水量と日照時間についても明確な関連は見取れませんでした。

* 他と離れて売れている月があります(2015年12月の2026個)





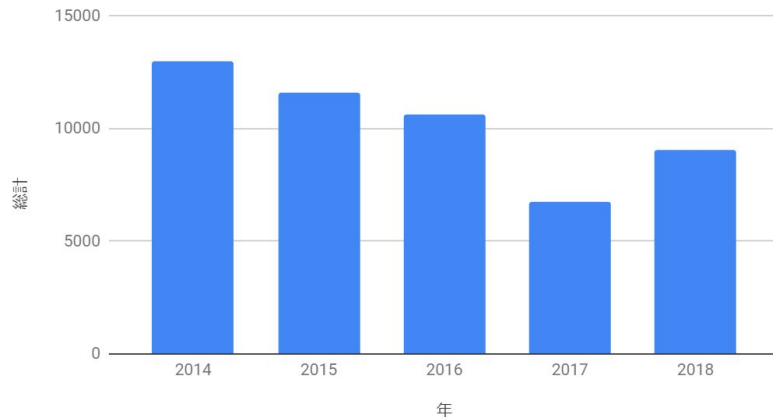
年間売上の傾向

全体として下がってましたが2018年に盛り返しました。

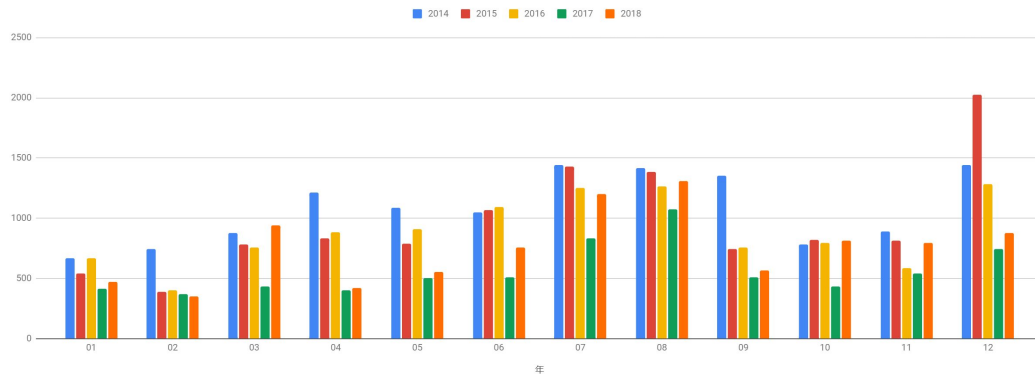
ほとんどの月で2017年の実績を2018年が上回りました。トレンドととらえるか、天候要素の異常なのかの判断が必要です。

*** 今回はそこまで盛り込めず**

年間売上総数の推移



2014、2015、2016、2017、2018





重回帰分析によるビールの需要予測式

説明変数として何を入れるか

→目に見えて売れる「7月・8月・12月」は必要性が高い

→天候とのデータはよくわからない&どの数値も連動度合いは変わらなそうだが、全く関係ないわけでもないと思われるので、一要素として「**平均気温**」を入れてみる

→2017年まで下がリトレンドだったのに、2018年で持ち直したことをどう判断するか→わからないけど「**前年同月**」を入れて見る

| 方針 | 変数 | 決定変数 | 出来上がったモデル |
|---------|------------------------------------------------|------------------------|---------------------------------------------------------------------------------------------------------------------------|
| フル | "7月フラグ", "8月フラグ", "12月フラグ", "平均気温(°C)", "前年同月" | 0.5938369939 556705 | 売上点数 = +310.245 * 7月フラグ +372.8618 * 8月フラグ +429.5901 * 12月フラグ +6.2224 * 平均気温(°C) +0.303 * 前年同月 +336.7826 |
| 前年同月を外す | "7月フラグ", "8月フラグ", "12月フラグ", "平均気温(°C)", | 0.5468324396 433233 | 売上点数 = +310.245 * 7月フラグ +372.8618 * 8月フラグ +429.5901 * 12月フラグ +6.2224 * 平均気温(°C) +0.303 * 前年同月 +336.7826 |
| 平均気温を外す | "7月フラグ", "8月フラグ", "12月フラグ", "前年同月" | 0.5876721382 113406 | 売上点数 = +354.6011 * 7月フラグ +418.2344 * 8月フラグ +363.4702 * 12月フラグ +0.3436 * 前年同月 +397.4833 |

| 方針 | 変数 | 決定変数 | 出来上がったモデル |
|-------------|--------------------------------------------------------------|--------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------|
| 月だけ | "7月フラグ", "8月フラグ", "12月フラグ" | 0.5177050597 389263 | 売上点数 = +528.1667 * 7月フラグ +606.9167 * 8月フラグ +583.4167 * 12月フラグ +649.3333 |
| 月を外す | "平均気温 (°C)", "前年同月" | 0.4745587927 1353674 | 売上点数 = +0.6081 * 前年同月 +6.0702 * 平均気温 (°C) +165.1579 |
| 最高気温を追加してみる | "7月フラグ", "8月フラグ", "12月フラグ", "平均気温 (°C)", "前年同月", "最高気温 (°C)" | 0.5940897850 363849 | 売上点数 = +309.897 * 7月フラグ +370.1044 * 8月フラグ +423.9553 * 12月フラグ +10.0711 * 平均気温 (°C) +0.3033 * 前年同月 -4.4503 * 最高気温 (°C) +384.5705 |

決定変数が上がらないのでデータを制限してみた

2017年まで減少トレンドだったのが2018年復活したことが予測の難易度を下げているのではと想定して、2018年のみの数値にしてみた。

結果決定変数は0.7を超えたが、反則的か？モデルの数値もこれまでと違う形式の様子..

| 方針 | 変数 | 決定変数 | 出来上がったモデル |
|------------------|-------------------------------------------------|------------------------|-----------------------------------------------------------------------------------------------------------------------------------------|
| 2018年だけのデータに限り分析 | "7月フラグ", "8月フラグ", "12月フラグ", "平均気温 (°C)", "前年同月" | 0.7402511461 282237 | 売上点数 = $-75.7825 * 7月フラグ$ $-423.2137 * 8月フラグ$ $-336.4274 * 12月フラグ$ $+1.9008 * 前年同月$ $-5.5461 * 平均気温 (°C)$ -156.8827 |



まとめ

- 7月・8月・12月の販売数が他の月に比べて高いことを織り込んだモデルとなりました。
- 気温や天候の影響は想定より低いようです。
- 今回の予測や切り口だけでは精度の高い需要予測を作ることは困難でした。さらなる予測、追加データなどを用いて深く分析する必要があります。