# Section 1

Attached is a database of a fictitious soft drink company. This is a time series data with vintage from 1/1/2012 to 1/9/2015 and a monthly granularity.

The data dictionary is as follows:

- Month – Month of activity

- SalesVol – Volume of Sales in litres

- TVGRP – TV Gross Rating Points – the number of times an ad is seen

- Instoreads – Spends on Ads inside the supermarket – posters,

signboards, danglers etc.

- Outdoorads – Spend on ads on billboards, hoardings, bus shelters etc.

- Promotions – Spends on promotional schemes like discounts, free

gift etc. targeting immediate sales

- Digitalads – Spends on media like Google Search ads / Google

Display ads / YouTube / FB / IG / Twitter etc.

- Price – Average price per 10 litres

- Comp1TV – Spends by 1st competitor on TV

- Comp1NP – Spends by 1st competitor on Newspaper

- Comp2OOH – Spends by 2nd competitor on outdoor advertising

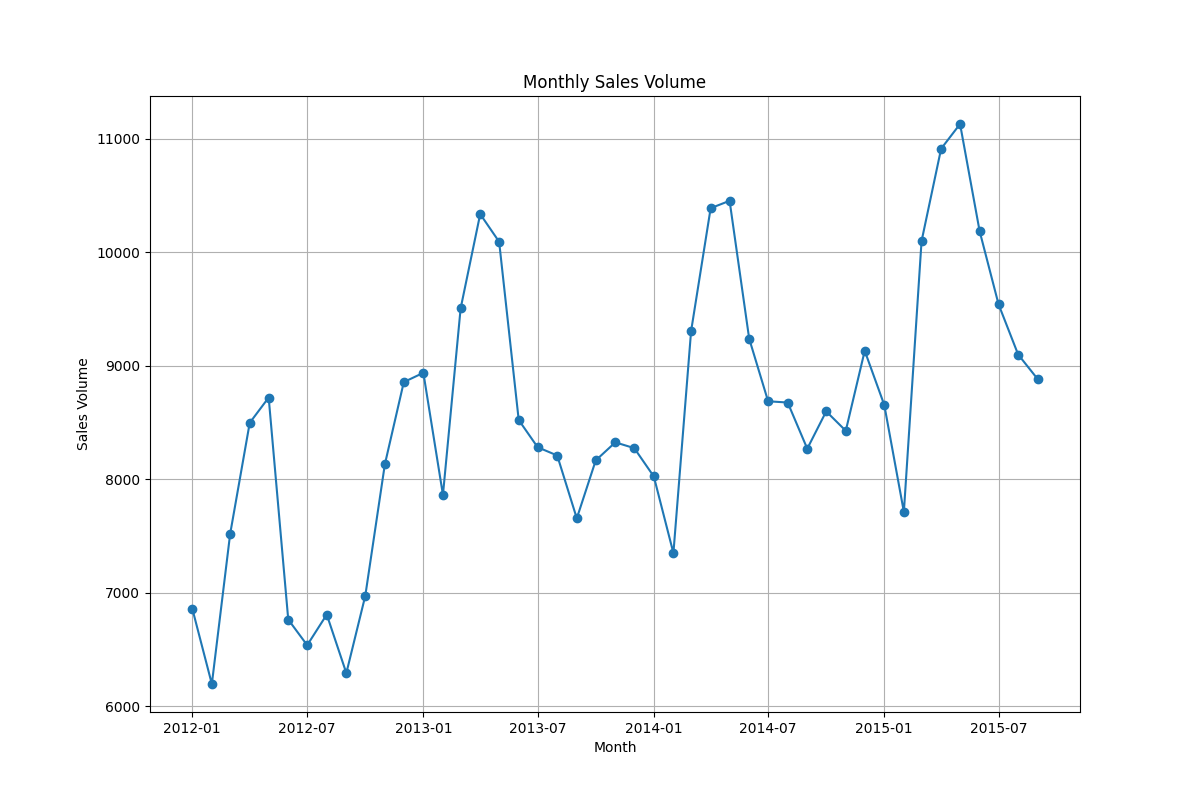
- Comp2NP – Spends by 2nd competitor on newspapers

Using your analysis skills, answer the following questions:

1. **Are soda sales afflicted by seasonality?**

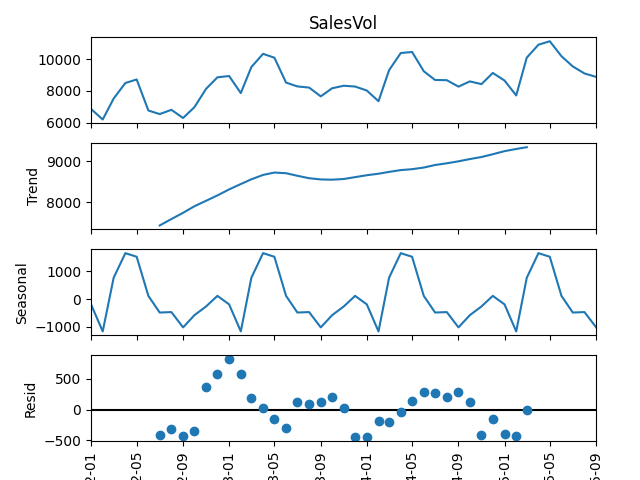
Yes, soda sales are afflicted by seasonality.

I have plotted the monthly SalesVol, where we can clearly see a repeating pattern (high and low occurring at regular interval) indicating seasonality in the data.



1. **Is there a trend seen on soda sales?**

Yes, there is a trend component in the seasonal decomposition as shown in the below figure.



1. **What are the correlations between volume sales & media advertising (TV/Instoreads/Outdoorads/Digital)?**

The below table shows the pearson correlation coefficient between the Sales Volume and different Media Advertising. The correlation is mild i.e., less than 50%. The maximum correlation is with OutdoorAds while the minimum is with DigitalAds.

| **Correlation** | **SalesVol** |
| --- | --- |
| **TVGrP** | 0.3421 |
| **InstoreAds** | 0.3510 |
| **OutdoorAds** | 0.4492 |
| **DigitalAds** | 0.0781 |

1. **What is the correlation between price and volume sales?**

The pearson correlation coefficient between “Volume Sales” and “Price” is approx. -0.45 indicating mild negative correlation. It shows the increase in “Price” leads to mild decrease in “Volume Sales” which seems intuitive as well.

1. **What is the correlation between sales volume and promotions?**

The correlation between “Volume Sales” and “Promotion” is approx. 0.15 indicating low positive correlation. It shows increase in “Promotions” leads to a small increase in “Volume Sales” but not as strong as media advertising.

1. **What is the partial correlation between sales and promotions, where TV/Instore/Outdoor/Digital have been partialled out?**

The partial correlation coefficient between sales and promotion is approx. 0.12 after removing the effect of media advertising. It indicates weak linear relationship even after removing the effect of media advertising.

1. **Which competitor and their media affect our sales the most?**

The below table shows the pearson correlation coefficient between SalesVolume and various Competitor’s media spend. The relationship is the strongest with the 2nd competitor’s outdoor advertising\*. The increase in 2nd competitor’s outdoor advertising leads to decrease in Sales Volume.

*(\*Please note that in the shared excel the column name is “Comp1OOH” however, I have treated it as “Comp2OOH as per the description”)*

| **Correlation** | **SalesVol** |
| --- | --- |
| **Comp1TV** | -0.2396 |
| **Comp1NPapers** | -0.0490 |
| **Comp2OOH** | -0.3405 |
| **Comp2NP** | -0.0910 |

TV GRP measures the number of times an ad is seen amongst an audience. It is a percentage and can exceed 100 as an ad can be seen multiple times by an individual.

Advertising has two properties:

1) Carry over effect

2) Saturation effect

Using Excel, transform the column TVGRP to create a new column 'Adstock' with the following formula:

A(t) = L × A(t−1) + g(t)^n / (g(t)^n + k^n)

A(t) is the adstock at time t, with the L × A(t−1) term being the carry over property.

g(t)^n / (g(t)^n + k^n) incorporates the saturation effect – sinusoidal curve.

Where:

- 0 ≤ L ≤ 1

- n > 0

- k > 0

**Please find attached the excel.**