Dirtea: Flavored cold tea

2025

A Guide to Consistent Brand Identity

MINDSET METRICS

PRASHANT ARYA
HENRI MAYOUD
FATIN MARWAT
RAGINI CHATTERJEE

STAY SAFE, DRINK DIRTEA.

AGENDA:

DESCRIPTIVE ANALYSIS AND INITIAL INSIGHTS:

- LOOKING INTO ALL THE MINDSET METRICS
- SEARCHING FOR CORELATION BETWEEN THE VARIABLES
- PLOTING THE TIME SERIES DATA

ATTITUDE METRICS

- CALCULATING POTENTIAL, STICKINESS AND RESPONSIVENESS FOR ALL THE MINDSET METRICS
- EXPLANATION OF TRANSFORMATIONS MADE TO THE DATA TO SIMPLIFY ANALYSIS

CONVERSION OF ATTITUDE METRICS

- CONSIDERATIONS FOR THE MULTIPLICATIVE MODEL FOR SALES REVENUE
- RESULTS

FURTHER ANALYSIS

- APPEAL
- SHORT-RUN GAIN
- LONG-RUN GAIN
- CONVERSION

WHAT IF ANALYSIS (ON THE DASHBOARD)

- CHANGE IN SALES GIVEN CHANGE IN MARKETING
- A LOOK INTO THE CHANGE IN MINDSET METRICS DUE TO CHANGE IN MARKETING TOOLS

EXPLORATION OF INTERACTION EFFECTS

- HOW MODELLING WITH INTERACTION EFFECTS CAN SHOW WHICH MARKETING MIX ITEMS ARE CODEPENDENT.
- HOW MINDSET METRICS ARE CODEPENDENT IN IMPACTS ON CONVERSION

PROJECT OUTLINE

ANALYSIS

- Descriptive analysis of all the variables
- Attitude metrics for the mindset metrics
- Further analysis

RESEARCH QUESTIONS:

- What can be inferred from the data?
- Which mindset metrics can be worked on to increase sales?
- Optimal mix of marketing tools to maximise sales?

DATA DESCRIPTION

SUMMARY STATISTICS:

Metric	Mean	SD	Min	Max	Median	NA Count
Sales	67505.68	3566.82	56117.41	78658.31	67521.78	0
Awareness	21.96	0.19	20.84	22.46	21.97	0
Liking	1.26	6.62	-19.72	22.83	1.64	0
Consideration	4.06	6.54	-16.50	25.65	4.30	0
InstagramAds	663.03	404.05	1.00	1499.00	638.50	0
TikTokAds	465.71	280.60	4.00	1000.00	450.00	0
SEA	1428.20	928.25	0.00	4491.00	1324.50	0
PoSPromotions	168.69	87.31	2.00	479.00	163.00	0
InfluencerColabs	798.81	153.16	379.00	1275.00	802.00	0

NOTES ON THE MINDSET METRICS

Awareness:
- Has low standard deviation as opposed to consideration and liking.

Liking:
- The negative minimum could represent instances of unfavorable customer sentiment.

Consideration:
- Negative values for consideration could indicate customers’ unwillingness to consider the product or service.

SUMMARY STATISTICS:

Metric	Mean	SD	Min	Max	Median	NA Count
Sales	67505.68	3566.82	56117.41	78658.31	67521.78	0
Awareness	21.96	0.19	20.84	22.46	21.97	0
Liking	1.26	6.62	-19.72	22.83	1.64	0
Consideration	4.06	6.54	-16.50	25.65	4.30	0
InstagramAds	663.03	404.05	1.00	1499.00	638.50	0
TikTokAds	465.71	280.60	4.00	1000.00	450.00	0
SEA	1428.20	928.25	0.00	4491.00	1324.50	0
PoSPromotions	168.69	87.31	2.00	479.00	163.00	0
InfluencerColabs	798.81	153.16	379.00	1275.00	802.00	0

NOTES ON THE CURRENT USE OF MARKETING TOOLS

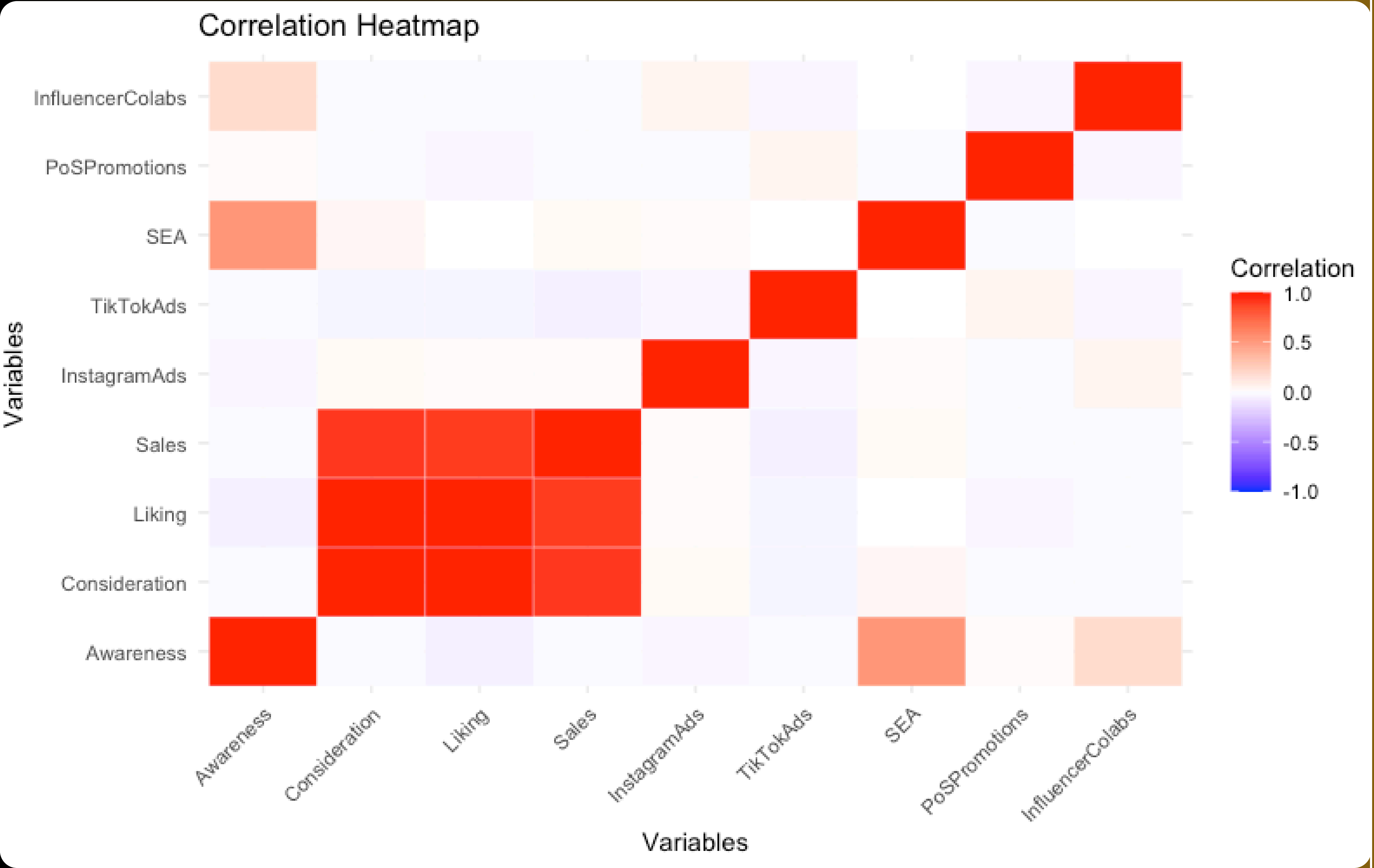
Instagram Ads:
- The mean ad spend on Instagram is ~663, with a wide range (1 to 1,499).
- High variability (SD = 404) suggests fluctuating investments in Instagram campaigns, possibly based on seasonal trends or experimental campaigns.

TikTok Ads:
- TikTok Ads have a mean spend of 465 with a range of 4 to 1,000.
- While the range is smaller compared to Instagram Ads, the standard deviation (~280) still reflects significant variability in spending.
-

Search Engine Ads (SEA):
- SEA spending shows the widest variability among all ad categories, with a range of 0 to 4,491 and a mean of 1,428.
- The high spend in SEA might indicate its importance in the current marketing strategy.

Point-of-Sale Promotions (PoSPromotions):
- The average spend on PoSPromotions is 168.7, with a range of 2 to 479.
- Lower variability (SD = 87.3) indicates more consistent spending patterns.

Influencer Collaborations:
- Influencer collaborations have an average spend of ~798, with a range of 379 to 1,275.
- This spending seems moderately consistent (SD = 153), suggesting a relatively stable collaboration strategy with influencers.



CORRELATION HEATMAP

INFERENCES

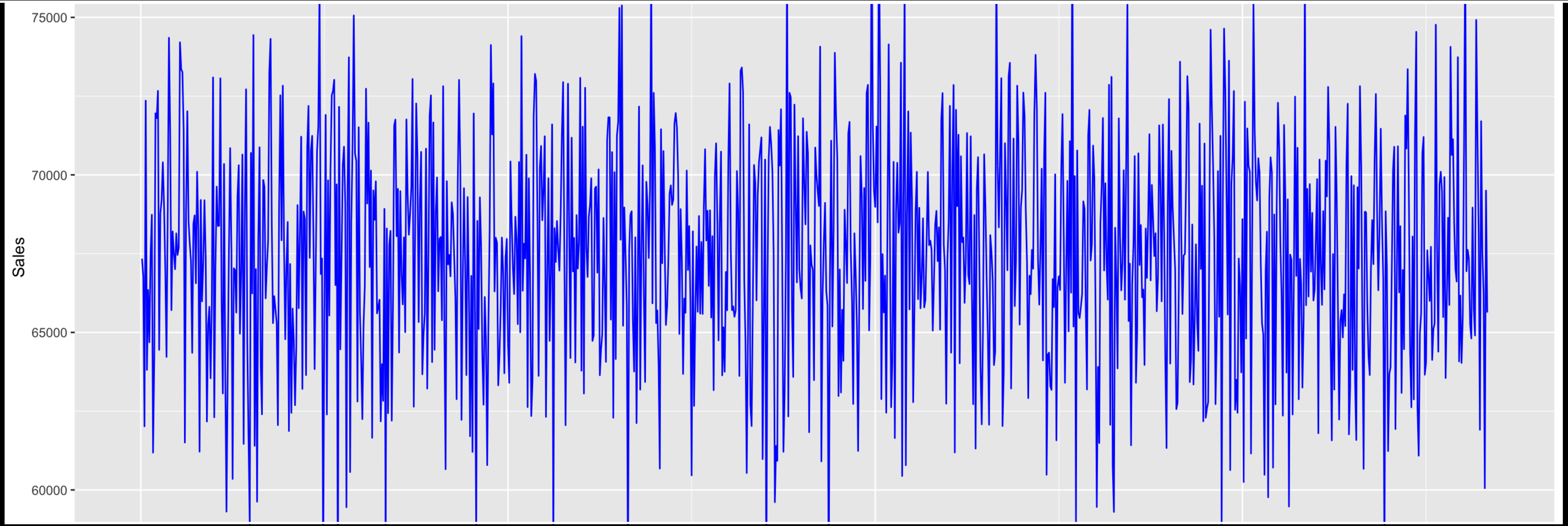
Sales:
Shows high corelation with Liking and Cosideration.
It shows mild corelation with TiktokAds.
Minor corelation with Awareness and PoSPromotions.

Awareness:
Shows mild correlation with SEA and InfluencerCollabs.

However, it is interesting to note that awareness is not highly correlated with sales. Additionally no other marketing tool is highly corealted with sales.

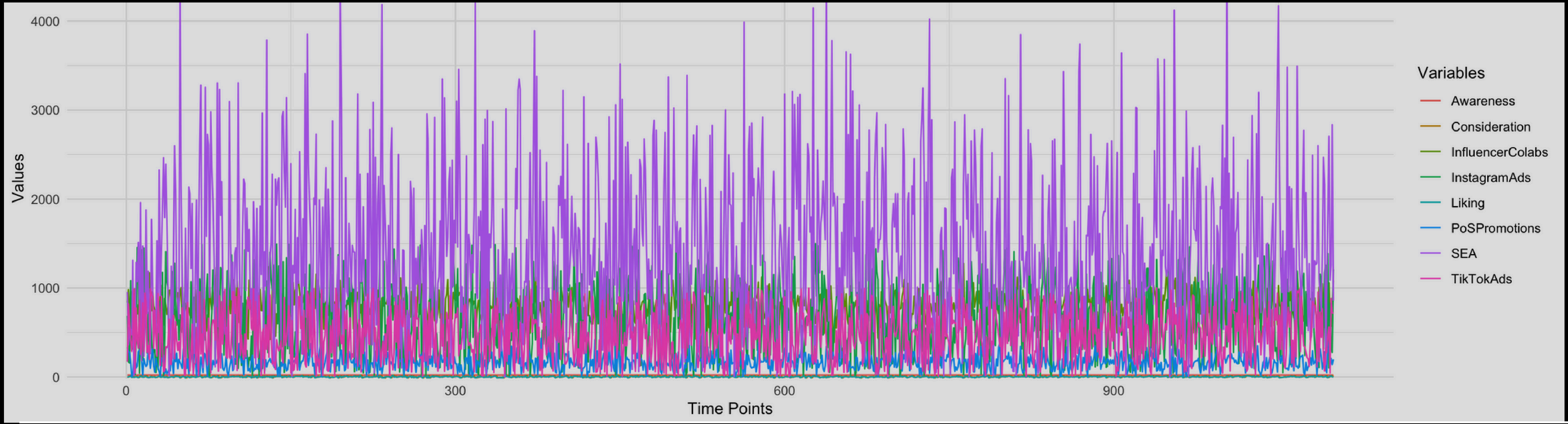
DIRTEA SALES AND OTHER METRICS OVER 1100 DAYS

2025



TIMESERIES VISUALISATION

- Sales time series (top) shows some seasonality and trend.
- The SEA in the bottom graph shows high variability.



METHOD DESCRIPTION

HERE WE TRANSFORMED THE DATA AND CALCULATED THE ATTITUDE METRICS (POTENTIAL, STICKINESS AND AWARENESS) TO MAKE INFERENCES ON THE MARKETING TOOLS WE SHOULD FOCUS ON

POTENTIAL:

Does attitude have room to grow?

$$\text{Potential\_Consideration} = \frac{100 - \text{mean}(df\$Consideration)}{100}$$

$$\text{Potential\_Awareness} = \frac{100 - \text{mean}(df\$Awareness)}{100}$$

$$\text{Potential\_Liking} = \frac{100 - \text{mean}(df\$Liking)}{100}$$

Potential	Value
Awareness	78.04%
Consideration	45.94%
Liking	48.74%

**As the range for liking and consideration ranges from -50 to 50 we have transformed the variables by adding 50 to each value so that they are all positive and easily useable for calculation of attitude metrics.*

INFERENCE

Awareness has the most potential followed by consideration and liking. Indicating that Awareness has the most room to grow.

STICKINESS: Does change in attitude stick?

Calculated by summing the coefficients of AR1 regressions for each attitude metric.

Stickiness	Value
Awareness	0.0974
Consideration	-0.1587
Liking	-0.1137

INFERENCE

Awareness: is the only mindset metric that has **positive stickiness.** This could indicate that people are able to recall the brand after initial exposure to it.

Consideration and Liking: Both consideration and liking have **negative stickiness**, which may indicate that peoples' opinions on the brand change after initial liking and consideration or misleading marketing.

RESPONSIVENESS: Can we move it?

LIKING

```
Call:
lm(formula = log(df$Liking + 1) ~ log(df$lag_liking + 1) + log(df$Instagram
Ads +
    1) + log(df$TikTokAds + 1) + log(df$SEA + 1) + log(df$PoSPromotions +
    1) + log(df$InfluencerColabs + 1), data = df)

Residuals:
     Min       1Q   Median       3Q      Max
-0.24160 -0.03751  0.00424  0.03974  2.18209

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                  1.9045903  0.1130502  16.847  < 2e-16 ***
log(df$lag_liking + 1)       0.5389515  0.0153959  35.006  < 2e-16 ***
log(df$InstagramAds + 1)     0.0010470  0.0027322   0.383  0.70165
log(df$TikTokAds + 1)       -0.0010863  0.0027877  -0.390  0.69686
log(df$SEA + 1)             -0.0073852  0.0026566  -2.780  0.00553 **
log(df$PoSPromotions + 1)    0.0004135  0.0037444   0.110  0.91209
log(df$InfluencerColabs + 1) -0.0050143  0.0133776  -0.375  0.70786
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

INFERENCE

From the regression analysis we can see that the only statistically significant coefficient on a marketing tool is for **search engine ads**.

However, the coefficient on SEA is a small negative number indicating that an increase in SEA may actually **decrease Liking**.

RESPONSIVENESS: Can we move it?

CONSIDERATION

```
Call:
lm(formula = log(df$Consideration + 1) ~ log(df$lag_consideration +
    1) + log(df$InstagramAds + 1) + log(df$TikTokAds + 1) + log(df$SEA +
    1) + log(df$PoSPromotions + 1) + log(df$InfluencerColabs +
    1), data = df)

Residuals:
     Min       1Q   Median       3Q      Max
-0.23641 -0.03783  0.00307  0.03904  2.04487

Coefficients:
                             Estimate Std. Error t value Pr(>|t|)
(Intercept)                 2.0497868  0.1107595  18.507   <2e-16 ***
log(df$lag_consideration + 1)  0.5018768  0.0154257  32.535   <2e-16 ***
log(df$InstagramAds + 1)    0.0013075  0.0026528   0.493    0.622
log(df$TikTokAds + 1)      -0.0009729  0.0027064  -0.359    0.719
log(df$SEA + 1)            -0.0042333  0.0025880  -1.636    0.102
log(df$PoSPromotions + 1)   0.0013694  0.0036345   0.377    0.706
log(df$InfluencerColabs + 1) -0.0052129  0.0129881  -0.401    0.688
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

INFERENCE

From the regression analysis we can see that none of the marketing tools have a statistically significant effect on consideration.

RESPONSIVENESS: Can we move it?

INFERENCE

How can we move Awareness:

From the regression analysis we can see that we can use **SEA (Search engine ads) and influencer Collabs** to move awareness.

As they are the only two independant varibales with statistically significant coefficients.

AWARENESS

```
    log(df$InstagramAds + 1) + log(df$TikTokAds + 1) + log(df$SEA +
    1) + log(df$PoSPromotions + 1) + log(df$InfluencerColabs +
    1), data = df)

Residuals:
       Min         1Q     Median         3Q        Max
-0.0205730 -0.0042111 -0.0000042  0.0042195  0.0187544

Coefficients:
                               Estimate Std. Error t value Pr(>|t|)
(Intercept)                   3.040e+00  9.086e-03 334.570  < 2e-16
log(df$lag_aware + 1)         5.093e-03  2.044e-03   2.492   0.0129
log(df$InstagramAds + 1)     -2.209e-04  1.926e-04  -1.147   0.2517
log(df$TikTokAds + 1)         6.845e-05  1.966e-04   0.348   0.7277
log(df$SEA + 1)               5.161e-03  1.896e-04  27.226  < 2e-16
log(df$PoSPromotions + 1)     1.447e-05  2.640e-04   0.055   0.9563
log(df$InfluencerColabs + 1)  6.449e-03  9.437e-04   6.834 1.37e-11

(Intercept)                  ***
log(df$lag_aware + 1)        *
log(df$InstagramAds + 1)
log(df$TikTokAds + 1)
log(df$SEA + 1)              ***
log(df$PoSPromotions + 1)
log(df$InfluencerColabs + 1) ***
```

RESPONSIVENESS: Can we move it?

SALES

INFERENCE

No marketing tool directly moves sales.

```
Call:
lm(formula = log(df$Sales + 1) ~ log(df$lag_sales + 1) + log(df$InstagramAds +
    1) + log(df$TikTokAds + 1) + log(df$SEA + 1) + log(df$PoSPromotions +
    1) + log(df$InfluencerColabs + 1), data = df)

Residuals:
      Min        1Q    Median        3Q       Max
-0.176247 -0.031930  0.002304  0.037513  0.272364

Coefficients:
                               Estimate Std. Error t value Pr(>|t|)
(Intercept)                  10.8571733  0.0759012 143.043  < 2e-16 ***
log(df$lag_sales + 1)         0.0242221  0.0047135   5.139 3.27e-07 ***
log(df$InstagramAds + 1)      0.0017549  0.0016020   1.095    0.274
log(df$TikTokAds + 1)        -0.0011596  0.0016349  -0.709    0.478
log(df$SEA + 1)               0.0007159  0.0015645   0.458    0.647
log(df$PoSPromotions + 1)    -0.0009417  0.0021961  -0.429    0.668
log(df$InfluencerColabs + 1) -0.0018097  0.0078473  -0.231    0.818
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

CONVERSION

```
lm(formula = log(df$Sales + 1) ~ log(df$lag_sales + 1) + log(df$Awareness +
    1) + log(df$Consideration + 1) + log(df$Liking + 1), data = df)

Residuals:
      Min        1Q    Median        3Q       Max
-0.076560 -0.014793  0.000388  0.014797  0.066005

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)               9.607916   0.267796  35.878  < 2e-16 ***
log(df$lag_sales + 1)     0.006591   0.002015   3.271  0.00111 **
log(df$Awareness + 1)    -0.041702   0.088144  -0.473  0.63623
log(df$Consideration + 1) 0.384455   0.084803   4.533 6.44e-06 ***
log(df$Liking + 1)        0.007573   0.079412   0.095  0.92404
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

INFERENCES

Consideration:

The only mindset metric with a statistically significant coefficient in the conversion regression.

From this regression we can also see that a **1% increase in consideration leads to a 0.38% increase in sales.**

APPEAL

Taking into account the Long Run impact and conversion the most appealing marketing elements from an investment point of view wold be **Instagram Ads** and **POS Promotions** (Point of sale promotions).

Platform	Value
Instagram	0.00021
TikTok	-0.00015
SEA	-0.00086
POS Promotion	0.00021
Collaborations	-0.00104

$$Appeal^{(i)} = \sum_{k=1}^{K} LR\ impact_k^{(i)} \times Conversion_k$$

Where long run impact is:

$$LR\ impact_k^{(i)} = Potential_k \times Responsiveness_k^{(i)} \times 1/(1 - Stickiness_k)$$

What-If Analyses and Long/Short-Run Gains

Assuming we wish to change our marketing mix in any way, we can use coefficients from responsiveness and conversion models to simulate various scenarios. An **interactive version of this is available on our Dashboard, allowing different marketing mixes.**

The gist of it is to set an amount we are increasing a specific method in the mix by, as a ratio, then multiply it by the respnsiveness associated with that marketing method.

$$New_{M_i} = Start_{M_i} * (Advertising_{new}/Advertising_{old})^{Responsiveness_{Advertising_{M_i}}})$$

Then we calculate short run gains in mindset metrics by taking these 3 values and measuring their change from the originals.

$$LRGain_{M_i} = Gain_{M_i}/(1 - Carryover_{M_i})$$

Subsequently, we calculate long run gains from these short run values by dividing them by the carryover (the coefficient of the lagged metric in the responsiveness models).

$$Contribution_{M_i} = LRGain_{M_i} * Conversion_{M_i}$$

Finally, we can calculate the contribution of the change in mindset metric to the gain (or loss) in sales by multiplying individual long run gains by the associated coefficients from the Conversion model.

Potential Interaction Effects?

```
> table_response_aware
                                          round.response_aware.coefficients..5.
(Intercept)                                                             3.02440
log(df[[lag_mindset]] + 1)                                              0.00540
log(df$InstagramAds + 1)                                               -0.00024
log(df$TikTokAds + 1)                                                   0.00009
log(df$SEA + 1)                                                         0.00726
log(df$PoSPromotions + 1)                                               0.00298
log(df$InfluencerColabs + 1)                                            0.00646
log(df$SEA + 1):log(df$PoSPromotions + 1)                              -0.00043
> table_response_consideration
                                          round.response_consideration.coefficients..5.
(Intercept)                                                                     2.32684
log(df[[lag_mindset]] + 1)                                                      0.50314
log(df$InstagramAds + 1)                                                       -0.04290
log(df$TikTokAds + 1)                                                          -0.00073
log(df$SEA + 1)                                                                -0.04402
log(df$PoSPromotions + 1)                                                       0.00184
log(df$InfluencerColabs + 1)                                                   -0.00657
log(df$InstagramAds + 1):log(df$SEA + 1)                                        0.00636
> table_response_liking
                                          round.response_liking.coefficients..5.
(Intercept)                                                              1.89556
log(df[[lag_mindset]] + 1)                                               0.54245
log(df$InstagramAds + 1)                                                -0.04995
log(df$TikTokAds + 1)                                                    0.02666
log(df$SEA + 1)                                                         -0.00691
log(df$PoSPromotions + 1)                                                0.03343
log(df$InfluencerColabs + 1)                                            -0.00636
log(df$InstagramAds + 1):log(df$SEA + 1)                                 0.00728
log(df$TikTokAds + 1):log(df$SEA + 1)                                   -0.00395
log(df$SEA + 1):log(df$PoSPromotions + 1)                               -0.00470
> table_response_sales
                                 round.response_sales.coefficients..5.
(Intercept)                                                   10.85717
log(df[[lag_mindset]] + 1)                                     0.02422
log(df$InstagramAds + 1)                                       0.00175
log(df$TikTokAds + 1)                                         -0.00116
log(df$SEA + 1)                                                0.00072
log(df$PoSPromotions + 1)                                     -0.00094
log(df$InfluencerColabs + 1)                                  -0.00181
```

if we design and test various models for responsiveness, we can check for interactions between various marketing mix elements. On the left are the final models we obtained after iteratively testing reduced models with less and less interactions modelled and comparing their Akaike Information Criteria (AIC).

We can see right away that responsiveness for sales has no notable interactions that improve the AIC criterion for useful information over the basic model.

In the cases of the mindset metrics, we can see a surprising interaction between Search Engine Advertising and Point of Sale Promotions, indicating when both increase together it has a dampening effect on the increased awareness.

For consideration, we can observe that there is a positive interaction effect for running more Instagram and search engine ads together, implying it slightly negates the negative effects of running each individually.

Liking Response has 3 observable interaction effects. It has the same Instagram-SEA interaction as consideration, but also shows that if Tiktok ads and SEA both increase together it adversely impacts Tiktok ads' positive impact. Additionally, increasing SEA and PoS Promotions together has increased negative impact, just as it does for Awareness, indicating it probably is not a good idea to invest in both together.

Potential Interaction Effects?

```
> table_conversion
                                                                                              round.conversion.coefficients..5.
(Intercept)                                                                                                        -14300.08370
log(lag_sales + 1)                                                                                                   1477.42763
log(df$Awareness + 1)                                                                                                4564.01200
log(df$consideration_transformed + 1)                                                                                3769.79150
log(df$liking_transformed + 1)                                                                                       -525.26971
log(lag_sales + 1):log(df$Awareness + 1)                                                                             -471.14884
log(lag_sales + 1):log(df$consideration_transformed + 1)                                                             -386.97089
log(lag_sales + 1):log(df$liking_transformed + 1)                                                                      -4.01526
log(df$Awareness + 1):log(df$consideration_transformed + 1)                                                         -1217.10706
log(df$Awareness + 1):log(df$liking_transformed + 1)                                                                  181.66747
log(df$consideration_transformed + 1):log(df$liking_transformed + 1)                                                  143.46551
log(lag_sales + 1):log(df$Awareness + 1):log(df$consideration_transformed + 1)                                        124.74429
log(df$Awareness + 1):log(df$consideration_transformed + 1):log(df$liking_transformed + 1)                            -45.73684
```

As on the previous slide, we can test various differing Conversion models using their AIC criterion to check for codependence and interactions between the the three mindset metrics as well as the previous day's sales.

This is very difficult and open-ended to interpret! We see a lot of surprises here as well. Awareness, Consideration and Liking increasing together has a very slight but negative impact, for instance. Conversely, if only 2 increase and one decreases, this term actually has a positive impact. First, it must be noted that liking on its own has a negative coefficient, but this does not necessarily mean that when liking increases the sales go down. For example, we notice a positive interaction between Liking and Awareness as well as liking and consideration. This means for liking to have an overall positive impact, Awareness or consideration need to increase more than liking by a factor of the sole liking coefficient and the interaction coefficients.

NOTE: These models with interactions were not used for our other calculation such as appeal or What-Ifs, as it is difficult and unituitive to generalize the models for this purpose.

REVIEWING RESEARCH QUESTIONS:

- **What can be inferred from the data?**

We can infer that most marketing tools do not seem to have a direct effect on sales. And we should look to influence sales through mindset metrics like consideration and liking.From our further analysis on the data we saw that consideration and liking are correlated with sales. However awareness is not.

- **Which mindset metrics can be worked on to increase sales?**

From our conversion regression analysis the only mindset metric that led to an increase in sales was consideration.

However, from our responsiveness regressions we see that the only mindset metric that can be positively moved with the marketing tools is awareness. Hence we use the dashboard to conduct a what if analysis and come up with an answer to our next question.

- **Optimal mix of marketing tools to maximise sales?**

The optimal mix of marketing tools to maximise sales would be to increase POS and instagram ads.

Thank you