MARKET *Mix Modeling*

Digital Cognition(R)
(https://github.com/digital-cognition-co-in)

TV ,
Website,
Facebook ...

Multi Channel Ad Spend

Available channels for Ad Spend …

Digital Cognition(R)

(https://github.com/digital-cognition-co-in)

Market ..

the action or business of promoting our products or services ...

Mix ..

..the types of promotion techniques and medium. TV , Social Media , YouTube , NewsPapers and Mags...

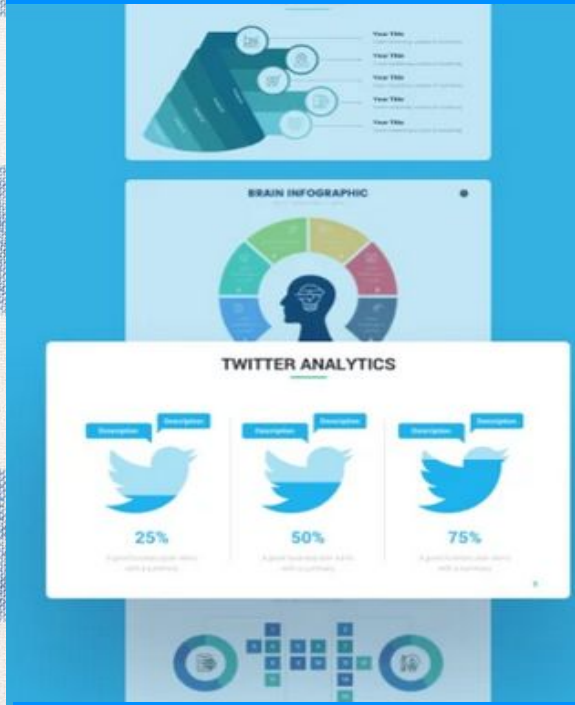
Modeling ..

relationships between various types of information (‘the mix part’) and inferring results .

The Statistical Tools

Why R and Python ?

The statistical tools of Choice - R and Pythoin

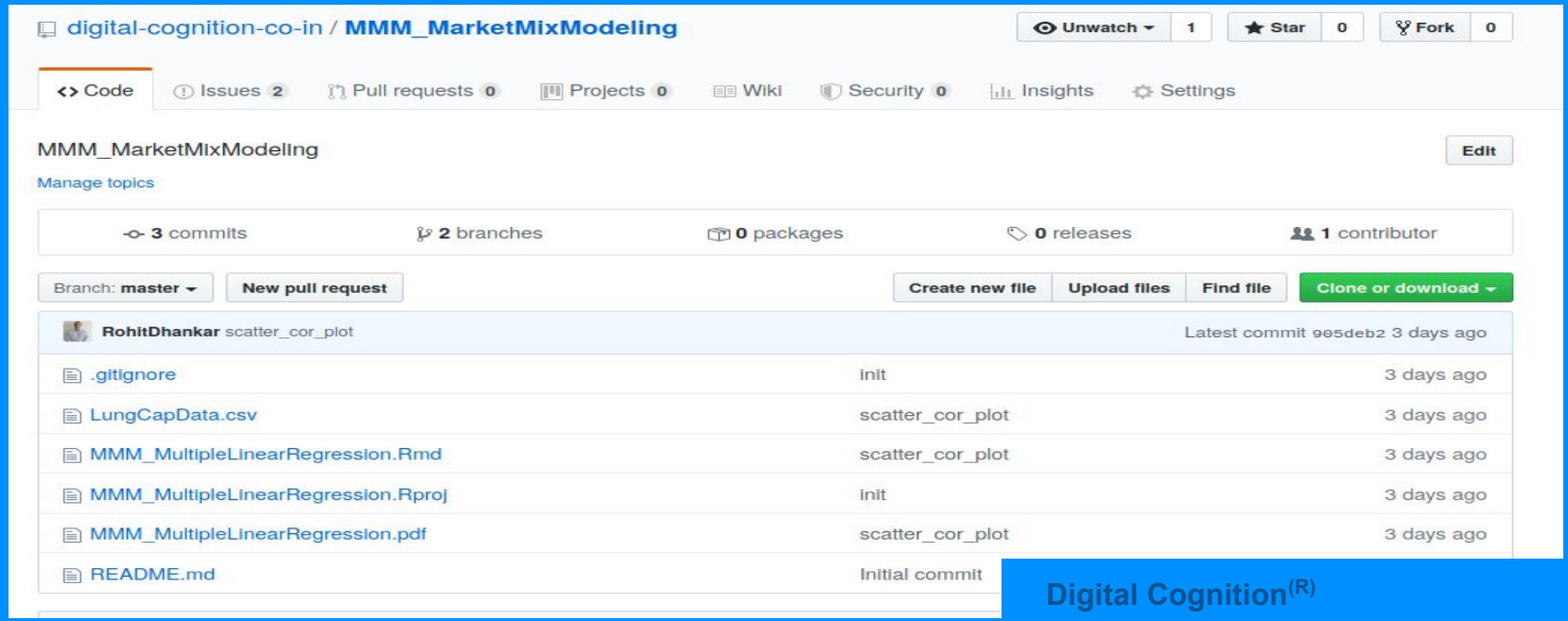


Digital Cognition(R)

(https://github.com/digital-cognition-co-in)

R & Python Code

https://github.com/digital-cognition-co-in



Digital Cognition(R)

(https://github.com/digital-cognition-co-in)

The Conceptual Framework

Why MMM ?

The Conceptual background for Market Mix Modeling

Digital Cognition(R)

(https://github.com/digital-cognition-co-in)

Linear Relationship - Dependent (Sales) & Independent Variables.

Failure to Isolate impact of Independent variables .

Non Linear Relationship - Decay of impact.

Linear Relation

Isolate Impact | ROI ambiguity | Non Linear Relation

Digital Cognition(R)

(https://github.com/digital-cognition-co-in)

MMM... as a statistical inference can simply be a case of **Multiple Linear Regression** or even advanced algorithms**...**

Estimate impact of various marketing tactics (marketing mix) on sales and predict / forecast - impact of future sets of tactics.

Used to ***optimize advertising mix*** and ***marketing budget spends***

Digital Cognition(R)
(https://github.com/digital-cognition-co-in)

Multiple Linear Regression

Sales (Dependent or Response Var.)

$$Sales_t = B_0 + B_1 Road..Hoard..Imp_t + B_2 TvAdSpend_t + B_p x_p + e_t$$

$Sales_t$ = Sales at a given time **- t ,** which is a Linear Function of the right side of the **Multiple Linear Regression Model** shown above.

Parameters $(B_0, B_1, B_{2....} B_p) B_1$, onwards for the ***p*** **(x_p)** independent variables.

epsilon (e_t) the Error term or Random term - this accounts for the variability in **$Sales_t$** which can not , be explained ,by the linear effect of the -*p* - independent variables.

- B_0 = Intercept on the Y Axis
- B_1 = Param for the RoadsideHoardingsImp..(Independent Var.) at time -t .
  It represents an ESTIMATE of Change in $Sales_t$ (Y-Dependent var.) corresponding to One Unit change in quant measure of - RoadsideHoardingsImpressions , when ***all other Independent Variables are held constant*** or are considered to be Constant.
- B_2 = Param for the TvAdSpend (Independent Var.)

Multiple Linear Regression

Sales (Dependent or Response Var.)

$$Sales_t = B_0 + B_1 Road..Hoard..Imp_t + B_2 TvAdSpend_t + B_p x_p + e_t$$

Multiple Linear Regression Model (shown above.)

$$E(y) = B_0 + B_1 x_1 + B_2 x_2 + \ldots + B_p x_p$$

Multiple Linear Regression Equation -

Assumed that ***mean or expected value*** of ***e (Epsilon)*** = 0

$$y\text{-}hat = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_p x_p$$

Estimated Multiple Linear Regression Equation -

where ($b_1 + b_2 + \ldots + b_p$) - are estimates of the population Params as population Params are usually not known .

Why - Multiple Linear Regression ?

Sales (Dependent or Response Var.)

How is a **Multiple Linear Regression Model** a better predictor of the Sales (Dependent or Response Var.) than maybe a **Simple Linear Regression Model** ?

The SSR and SSE Values will change when we add - Multiple Independent variables , thus we infer that a **Multiple Linear Regression Model** is a better predictor of the Sales.

Relation between the SST , SSR and SSE :-

$$SST = SSR + SSE$$

$$SST = \text{total sum of squares} = \Sigma(y_i - ybar)^2$$

$$SSR = \text{sum of squares due to regression} = \Sigma(yhat_i - ybar)^2$$

$$SSE = \text{sum of squares due to error} = \Sigma(y_i - yhat_i)^2$$

Why - Multiple Linear Regression ?

Sales (Dependent or Response Var.)

Multiple **Coefficient Of Determination (R^2 R-squared)**

$$R^2 = SSR/SST$$

Proportion of variability in the **dependent variable** which is **explained** by the **Estimated multiple regression equation.** Its generally observed that **R-squared will increase as we add variables to a - multiple regression equation.**

Thus for MMM we will get a Higher value for R^2 with - y ~ RoadsideHoardingsImp + TvAdSpend + x_p

Vs.

y ~ RoadsideHoardingsImp ….only

Adjusted Multiple Coefficient Of Determination ($R^2_{(adj)}$ Adjusted-R-squared)

$$R^2_{adj} = 1-(1-R^2)n-1/n-p-1$$

n = count of Obs.
p = count of independent var.

F-test **Test for Overall Significance**

Within the scope of **Multiple Linear regression - the F Test** is called the **- Test for Overall Significance ,** as it determines if a **Significant Relationship** exists between the **Dependent Var and All the Independent Var's .** After the F-test is used to **- Test for Overall Significance ,** a separate **t-test** is conducted for all independent variables.

$H_0 = B_0 = B_1 = B_2 = B_p = 0$
This Null Hypothesis - states that all PARAMS are Equal to Zero

H_a
The alternative Hypothesis - states that atleast one PARAM is Not Equal to Zero
We want to Reject the Null Hypothesis and accept the alternative.

This leads us to **MEAN SQUARE ERRORS** - the **MSE** and the **MSR**

F-test **Test for Overall Significance**

Within the scope of **Multiple Linear regression - the F Test** is called the **- Test for Overall Significance ,** as it determines if a **Significant Relationship** exists between the **Dependent Var and All the Independent Var's .** After the F-test is used to **- Test for Overall Significance ,** a separate **t-test** is conducted for all independent variables.

$H_0 = B_0 = B_1 = B_2 = B_p = 0$
This Null Hypothesis - states that all PARAMS are Equal to Zero

H_a
The alternative Hypothesis - states that atleast one PARAM is Not Equal to Zero
We want to Reject the Null Hypothesis and accept the alternative.

This leads us to **MEAN SQUARE ERRORS** - the **MSE** and the **MSR**

Our References :-

OLS - Ordinary Least Squares = https://en.wikipedia.org/wiki/Ordinary_least_squares

from statsmodels.formula.api import ols =
https://scipy-lectures.org/packages/statistics/index.html#multiple-regression-including-multiple-factors

ISLR - An Introduction to Statistical Learning - with Applications in R -
http://faculty.marshall.usc.edu/gareth-james/ISL/

ggPredict() - Visualize multiple regression model

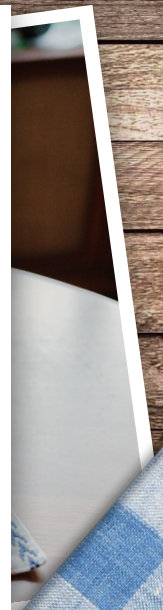
https://cran.r-project.org/web/packages/ggiraphExtra/vignettes/ggPredict.html

Cornell Univ -
http://mezeylab.cb.bscb.cornell.edu/labmembers/documents/supplement%205%20-%20multiple%20regression.pdf

Our References :-

Media TRP - https://marketing-dictionary.org/t/target-rating-point/

Gross Rating Points or GRPs -
https://www.srds.com/frontMatter/sup_serv/calculator/grp_trp/grps_trps.html
https://mpwmarketing.com/2017/05/efficiently-media-buy-tv/



Digital Cognition(R)

Code Repository

Python Code:

Digital Cognition(R)
(https://github.com/digital-cognition-co-in)

Digital Cognition(R)

Code Repository

R Code:

Digital Cognition(R)

QnA

Rohan Mathur:

https://www.linkedin.com/in/rohanmathur17/

Rohit Dhankar:

https://www.linkedin.com/in/rohitdhankar/