**SQL QUERY FOR CORRELATION**

SELECT \*

FROM `inbound-entity-382011.el\_vh.ele\_vhl`

#------Corrlation Test--------#

SELECT factor\_encourage,

       factor\_discourage

FROM `inbound-entity-382011.el\_vh.ele\_vhl`

SELECT CORR(factor\_encourage, factor\_discourage) AS correlation

FROM  `inbound-entity-382011.el\_vh.ele\_vhl`

**SQL QUERY FOR CHI SQUARE ONE**

SELECT \* FROM pro.electrical\_vehicles;

USE pro;

CREATE TABLE contingency\_tables

SELECT occupation,

prefer\_ev\_or\_fv,

COUNT(\*) AS counter

FROM pro.electrical\_vehicles

GROUP BY occupation,

prefer\_ev\_or\_fv;

SELECT \*

FROM contingency\_tables

WITH

contingency\_tables AS ( SELECT occupation, prefer\_ev\_or\_fv,

COUNT(\*) AS counter

FROM contingency\_tables

GROUP BY occupation, prefer\_ev\_or\_fv

),

table\_totals AS

( SELECT occupation,

SUM(counter) AS row\_total

FROM contingency\_tables

GROUP BY occupation

),

grand\_total AS

( SELECT SUM(counter) AS total

FROM contingency\_tables )

SELECT contingency\_tables.occupation, contingency\_tables.prefer\_ev\_or\_fv,

(table\_totals.row\_total \* SUM(contingency\_tables.counter)) / grand\_total.total AS expected\_value, contingency\_tables.counter

FROM contingency\_tables

JOIN table\_totals ON contingency\_tables.occupation = table\_totals.occupation

CROSS JOIN grand\_total;

WITH

contingency\_tables AS

( SELECT occupation, prefer\_ev\_or\_fv,

COUNT(\*) AS counter

FROM contingency\_tables

GROUP BY occupation, prefer\_ev\_or\_fv

),

table\_totals AS

( SELECT occupation,

SUM(counter) AS row\_total

FROM contingency\_tables

GROUP BY occupation

),

grand\_total AS

( SELECT SUM(counter) AS total

FROM contingency\_tables

),

expected\_values AS

( SELECT contingency\_tables.occupation, contingency\_tables.prefer\_ev\_or\_fv,

(table\_totals.row\_total \* SUM(contingency\_tables.counter)) / grand\_total.total AS expected\_value, contingency\_tables.counter AS observed\_value

FROM contingency\_tables

JOIN table\_totals ON contingency\_tables.occupation = table\_totals.occupation

CROSS JOIN grand\_total )

SELECT SUM((observed\_value - expected\_value) \* (observed\_value - expected\_value) / expected\_value) AS chi\_square

FROM expected\_values;

**SQL QUERY FOR CHI SQUARE TWO**

SELECT \* FROM pro.electrical\_vehicles;

CREATE TABLE contingency\_table\_two

SELECT education\_level,

knowledge\_ev,

COUNT(\*) AS counter

FROM pro.electrical\_vehicles

GROUP BY education\_level,

knowledge\_ev;

SELECT \*

FROM contingency\_table\_two

WITH

contingency\_table\_two AS

( SELECT education\_level, knowledge\_ev,

COUNT(\*) AS counter

FROM contingency\_table\_two

GROUP BY education\_level, knowledge\_ev

),

table\_totals AS

( SELECT education\_level,

SUM(counter) AS row\_total

FROM contingency\_table\_two

GROUP BY education\_level

),

grand\_total AS

( SELECT SUM(counter) AS total

FROM contingency\_table\_two)

SELECT contingency\_table\_two.education\_level, contingency\_table\_two.knowledge\_ev,

(table\_totals.row\_total \* SUM(contingency\_table\_two.counter)) / grand\_total.total AS expected\_value, contingency\_table\_two.counter

FROM contingency\_table\_two

JOIN table\_totals ON contingency\_table\_two.education\_level = table\_totals.education\_level

CROSS JOIN grand\_total;

WITH

contingency\_table\_two AS

( SELECT education\_level, knowledge\_ev,

COUNT(\*) AS counter

FROM contingency\_table\_two

GROUP BY education\_level, knowledge\_ev

),

table\_totals AS

( SELECT education\_level,

SUM(counter) AS row\_total

FROM contingency\_table\_two

GROUP BY education\_level

),

grand\_total AS

( SELECT SUM(counter) AS total

FROM contingency\_table\_two

),

expected\_values AS

( SELECT contingency\_table\_two.education\_level, contingency\_table\_two.knowledge\_ev,

(table\_totals.row\_total \* SUM(contingency\_table\_two.counter)) / grand\_total.total AS expected\_value, contingency\_table\_two.counter AS observed\_value

FROM contingency\_table\_two

JOIN table\_totals ON contingency\_table\_two.education\_level = table\_totals.education\_level

CROSS JOIN grand\_total ) SELECT SUM((observed\_value - expected\_value) \* (observed\_value - expected\_value) / expected\_value) AS chi\_square FROM expected\_values;

**ANOVA**

values <- c(222,82,39,14,153,121,59,24,146,103,52,56,125,118,41,73,133,103,37,84,142,113,58,44,159,99,54,45,178,110,46,23,168,106,53,30,154,118,46,39)

types <- c(rep("Tesla", 4), rep("BMW i-series", 4), rep("Renault ZOE", 4),

+ rep("Chevrolet volt", 4),

+ rep("Cadillac Poolside", 4),

+ rep("Toyota Prius", 4),

+ rep("Hyundai Kona", 4),

+ rep("Tata Nexon", 4),

+ rep("Mahindra", 4),

+ rep("MG", 4))

features <- rep(c("Deep knowledge of features", "Superficial awareness", "Sounds familiar", "Never heard of this model"), 10)

av <- data.frame(features, types, values)

anov <- aov(values ~ types + features, data = av)

summary(anov)

library(RMySQL)

> con <- dbConnect(MySQL(), user = "root", password = "1303", dbname = "pro", host = "localhost")

> anvt <- summary(aov(anov))

> anovatw <- data.frame(anvt$table)

> dat <- data.frame(summary(anov))