In this Project :

1. I selected all the data and filtered if there is any blank cell or not

2. Checked and removed duplicates

A screenshot of a computer

Description automatically generated

3. Found Independent and dependent variables

Independent Variables = size and total\_bill

Dependent variable = Tips

4. Used Multiple regression analysis to obtain data model as shown in screenshots and excel files

A screenshot of a computer

Description automatically generated

A screenshot of a document

Description automatically generated

5. Converted categorical values into numerical using IFS condition

Using formula =IFS(tips!$A2="male",0,tips!$A2="female",1)

=IFS(tips!$B2="No",0,tips!$B2="yes",1)

=IFS(tips!$C2="thur",0,tips!$C2="fri",1,tips!$C2="Sat",2,tips!$C2="Sun",3)

=IFS(tips!$D2="Dinner",0,tips!$D2="Lunch",1)

A screenshot of a computer

Description automatically generated

5. Then using predictive analysis multiple regression formula, found the predictive tips along side actual ones

Using formula =$Q$20+($E2\*$Q$21)+($F2\*$Q$22)

6. Then found errors in next column

Using formula =$H2-$G2

A screenshot of a table with numbers

Description automatically generated

7. Calculated RMSE

Using Formula =L4/L5

A screenshot of a graph

Description automatically generated

Calculated Correlation Coefficient

Using Formula =CORREL(E1:E244,G1:G244)

=CORREL(F1:F244,G1:G244)

A green rectangle with black text

Description automatically generated

Thank you