Below Procedures were run for the wine data set analysis:

|  |  |
| --- | --- |
| **SAS PROGRAM** | **Comments** |
| %let path=/folders/myfolders/Popularity;  libname news "&path"; | Set up the wine library |
| ods graphics / imagemap=on;  proc reg data=news.popularity plots(only)=(cp);  ALL\_REG: model \_shares=  \_kw\_avg\_avg  \_num\_hrefs  \_num\_imgs  \_self\_reference\_avg\_sharess  \_title\_subjectivity  \_num\_keywords  \_num\_videos  \_avg\_negative\_polarity  \_avg\_positive\_polarity  \_is\_weekend  \_rate\_negative\_words  \_rate\_positive\_words  /selection = cp rsquare adjrsq best=10;  title 'Best Models Using All-Regression Option';  run;  quit;  title; | All regression approach |
| proc glm data=news.popularity  plots(only)=(contourfit);  model \_shares=\_kw\_avg\_avg \_self\_reference\_avg\_sharess;  store out=multiple;  title 'Model with Avg Keyword and Avg Self reference shares';  run;  quit; | Multivariate regression |
| proc reg data=news.popularity;  model \_shares = \_kw\_avg\_avg;  title 'Popularity Prediction Using Average Keywords';  run;  quit;  title; | Linear regression |
| ods graphics/imagemap=on;  proc reg data=news.popularity plots(only)=(cp);  predict: model \_shares=  \_kw\_avg\_avg  \_num\_hrefs  \_num\_imgs  \_self\_reference\_avg\_sharess  \_title\_subjectivity  \_num\_keywords  \_num\_videos  \_avg\_negative\_polarity  \_avg\_positive\_polarity  \_is\_weekend  \_rate\_negative\_words  \_rate\_positive\_words;  explain: model \_shares=  \_kw\_avg\_avg  \_num\_hrefs  \_num\_imgs  \_self\_reference\_avg\_sharess  \_title\_subjectivity  \_num\_keywords  \_num\_videos  \_avg\_negative\_polarity  \_avg\_positive\_polarity  \_is\_weekend  \_rate\_negative\_words  \_rate\_positive\_words;  title'Best Model for Prediction';  run;  quit; | Best model for prediction |
| proc sql;  create table Category1 as  select count(\_data\_channel\_is\_bus) as Business  from news.popularity where \_data\_channel\_is\_bus=1;  quit;  proc sql;  create table Category2 as  select count(\_data\_channel\_is\_entertainment) as Entertainment  from news.popularity where \_data\_channel\_is\_entertainment=1;  quit;  proc sql;  create table Category3 as  select count(\_data\_channel\_is\_lifestyle) as Lifestyle  from news.popularity where \_data\_channel\_is\_lifestyle=1;  quit;  proc sql;  create table Category4 as  select count(\_data\_channel\_is\_socmed) as SocialMedia  from news.popularity where \_data\_channel\_is\_socmed=1;  quit;  proc sql;  create table Category5 as  select count(\_data\_channel\_is\_tech) as Technology  from news.popularity where \_data\_channel\_is\_tech=1;  quit;  proc sql;  create table Category6 as  select count(\_data\_channel\_is\_world) as World  from news.popularity where \_data\_channel\_is\_world=1;  quit; | Creating tables for different categories. |
| proc transpose data=category1  out=CatBUS;  run;  proc transpose data=category2  out=CatENT;  run;  proc transpose data=category3  out=CatLIF;  run;  proc transpose data=category4  out=CatSOC;  run;  proc transpose data=category5  out=CatTECH;  run;  proc transpose data=category6  out=CatWOR;  run; | Transposing the table |
| Proc contents data=news,popularity;  Run; | Looking into contents of dataset |
|  |  |