Title: Literature Review of Correlation matrix

Authors : Michael Friendly

Source : http://www.datavis.ca/papers/corrgram.pdf

What does the paper talk about?

This paper describes a set of techniques we subsume under the name “corrgram”, based on the rendering the value of a correlation to depict its sign and magnitude.

What does it relate to?(Technique we plan to use)

A correlation matrix is a table showing correlation coefficients between variables. Each cell in the table shows the correlation between two variables. A correlation matrix is used to summarize data, as an input into a more advanced analysis, and as a diagnostic for advanced analyses.

What are the advantages of Regression Analysis?

1. To summarize a large amount of data where the goal is to see patterns. In our example above, the observable pattern is that all the variables highly correlate with each other.
2. To input into other analyses
3. As a diagnostic when checking other analyses. For example, with linear regression, a high amount of correlations suggests that the linear regression estimates will be unreliable.

Any assumptions in the paper?

No assumptions made.

What are the main claims?

The correlation matrix plays an important role in multivariate analysis since by itself it captures the pairwise degrees of relationship between different components of a random vector.
We can claim to have presented a more general and comprehensive account of the possibilities than has appeared previously. We have also (a) suggested a new scheme for ordering variables in such displays, (b) extended the idea of correlation mapping to more general concepts of dependence and independence, and (c) illustrated (we hope convincingly) why they might be useful.

What is the takeaway from this paper for you?

This paper provides information about correlation matrix and how good we can use it to visualize the data and their relation with other data and use it with our dataset.

Title: Review of Regression Analysis Models

Authors- Dr. Amita Goel, Aviral Gupta, and Akshay Sharma

Source- Maharaja Agrasen Institute of Technology New Delhi, India.

https://www.ijert.org/research/review-of-regression-analysis-models-IJERTV6IS080060.pdf

Year- 2017

What does it relate to? (Technique we plan to use)

Regression Analysis is a technique to find out the relationship between different variables. Regression looks closely into how a dependent variable is affected upon varying an independent variable while keeping the other independent variables constant.

What are the advantages of Regression Analysis?

1. By using the relevant model to a data set, Regression Analysis can accurately predict a lot of useful information like Stock Prices, Medical Conditions and even Sentiments of the public.
2. Results from regression analysis adds a scientific backing to a decision or policy and makes it even more reliable as it likelihood of success is then high.
3. Sometimes, an anomaly between the prediction of regression analysis and a decision/thinking can help correct the fallacy of the decision.
4. Large data sets realize their potential to provide new dimensions to a study through the application of Regression Analysis.

Any assumptions in the paper?

No assumptions

What are the main claims?

Regression is a really effective tool in statistical analysis of data. Different Regression models are used in different situations. A Linear Regression Model is best suited to data representing a linear relationship between two variables, a Polynomial Regression Model is used in case of multiple variables having a polynomial relationship.

What is the takeaway from this paper for you?

This paper provides a good insight on linear regression and multiple regression models that can be used in our project.

Title: Data Analysis on 'Nutrition Facts for McDonald's Menu' Data-set using Python

Author: Neha Tiwari , Prof. Vaishali Gatty

Source: Vivekanand Education Society's Institute of Technology, India

https://ijesc.org/upload/c16ba3a5d222cfa2452ab986ec76af5a.Data%20Analysis%20on%20Nutrition%20Facts%20for%20McDonalds%20Menu%20Data%20set%20using%20Python.pdf

Year: 2017

What does it relate to?

In this paper, the data analysis of nutritious and non-nutritious food items in 'Nutrition Facts for McDonald's Menu' data-set is done using Python language.

Any Assumptions?

No assumptions

What are the main claims?

It is analyzed that the items in menu dataset can be categorized as nutritious food and non-nutritious food based on different chart diagrams and range values in percentage. It is beneficial for demonstrating different range values for food nutrients such as vitamin A, vitamin B, vitamin C, sugar, dietary fibers, fats, carbohydrates, cholesterol, iron, sodium and protein for their proper consumption from menu items.

What is the takeaway from this paper for you?

This paper provides a good insight on what features can be useful to design a model that can be used in our project.