Jake Simpkins  
jakesbusinessacc@outlook.com

Top of Form

Bottom of Form

R studio & Power BI

Just IT Assignment

Interview questions   
  
What is the difference between Univariate, bivariate and multivariate?

Univariate statistics summarise only one variable at a time. Bivariate statistics compare two variables. Multivariate statistics compare more than two variables.  
  
What is an outlier and how to identify them?  
Outliers are extreme values that differ from most other data points in a dataset.

...

There are four ways to identify outliers:

Sorting method.

Data visualisation method.

Statistical tests (z scores)

Interquartile range method.  
  
What is correlation?  
Quick definition: Correlation analysis, also known as bivariate, is primarily concerned with finding out whether a relationship exists between variables and then determining the magnitude and action of that relationship. Key takeaways: Correlation does not equal causation.   
  
  
R-Studio Screenshots  
  
PATH, Load data, view data and install. Packages  
  
Graphical user interface, text, application

Description automatically generated

Checking the data frame, checking for missing values, drop missing values, making sure rows are removed , checking for duplicates, round enough values to two places   
  
  
  
Graphical user interface, text, application

Description automatically generated  
  
Viewing the data frame , checking the outliers using boxplot, create a boxplot that highlight the outliers, remove outliers in profitability

Graphical user interface, application, Word

Description automatically generated  
  
Graphical user interface, application

Description automatically generated

Remove outliers from Worldwide gross  
  
Graphical user interface, text, application

Description automatically generated  
  
View data frame , Summary statistics Univariate Analysis, Bivariate and Scatterplot, create a barchart and export clean data.  
Graphical user interface, application

Description automatically generated  
Graphical user interface, application

Description automatically generated

Power BI visuals

Average of Rotten tomatoes by Genre   
  
Graphical user interface, application

Description automatically generated  
  
Sum of Audience score by film   
  
Graphical user interface, application, table, Excel

Description automatically generated

Sum of Worldwide gross by Genre   
Graphical user interface

Description automatically generated  
  
Count of film by year  
A screenshot of a computer

Description automatically generated  
  
Sum of profitability by lead.studioChart

Description automatically generated