Homework Assignment-4

Avisha Singh

Section-A

Ans1- A dendrogram is used to show the hierarchical relationship between objects and is the output of the hierarchical clustering. A dendrogram could potentially help with identifying the number of clusters to choose when applying hierarchical clustering.

It is used to work out the best way to allocate objects to clusters.

Ans2- Autoregressive Integrated Moving Average (ARIMA)-

It combines different autoregressive models with moving average. The auto regressive part is used to show that it is trained on its past data points. The Autoregressive Integrated Moving Average (ARIMA) method models the next step in the sequence as a linear function of the previous time steps' differenced observations and residual errors.

It combines both Autoregression (AR) and Moving Average (MA) models, as well as a sequence differencing pre-processing step called integration, to make the sequence stationary (I).

The model's notation entails specifying the order of the AR(p), I(d), and MA(q) models as parameters to an ARIMA function, e.g. ARIMA (p, d, q). ARIMA models can be used to create AR, MA, and ARMA models.

Vector Autoregression (VAR)-

Using an AR model, the Vector Autoregression (VAR) approach models the next step in each time series. It's when AR is used to numerous parallel time series, such as multivariate time series.

The model's nomenclature entails expressing the AR(p) model's order as parameters to a VAR function, such as VAR (p).

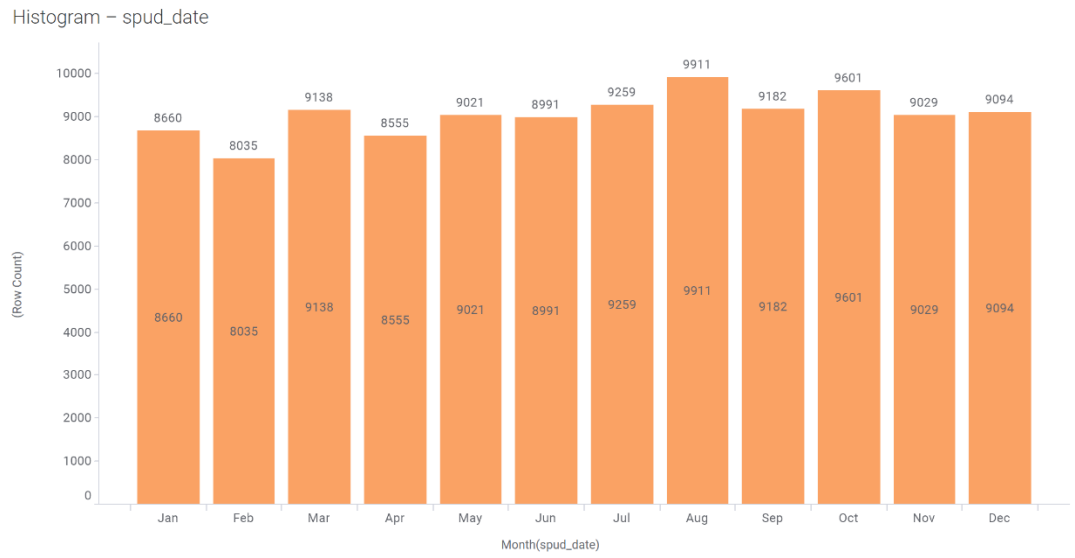
The approach works well with multivariate time series that don't have any trend or seasonal components.

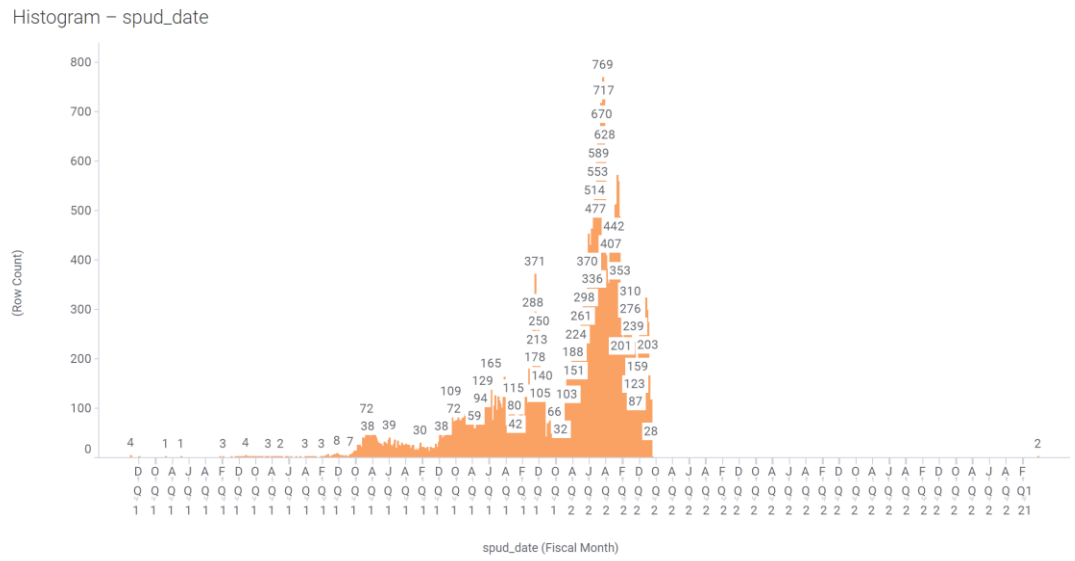
Ans3- a-Select public data source (GWPA, Colorado, North Dakota) and extract (as excel file) a well data set of at least 100 wells and all additional related information.

I have selected public data source (GWPA, Colorado, North Dakota) and I have extracted (as excel file) a well data set of at least 100 wells and all additional related information.

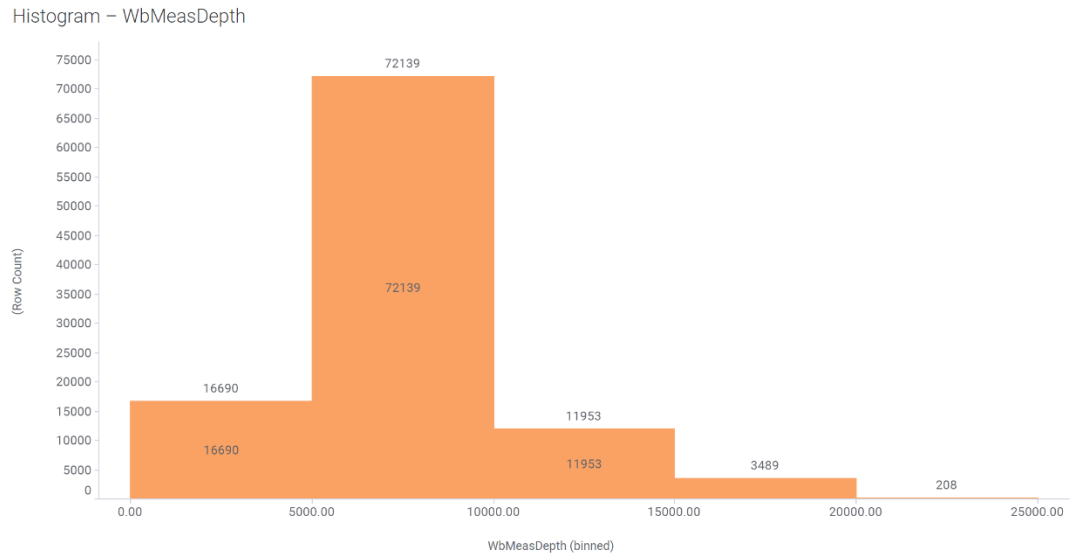
b. Plot number of wells against spud date (use Spotfire)

-

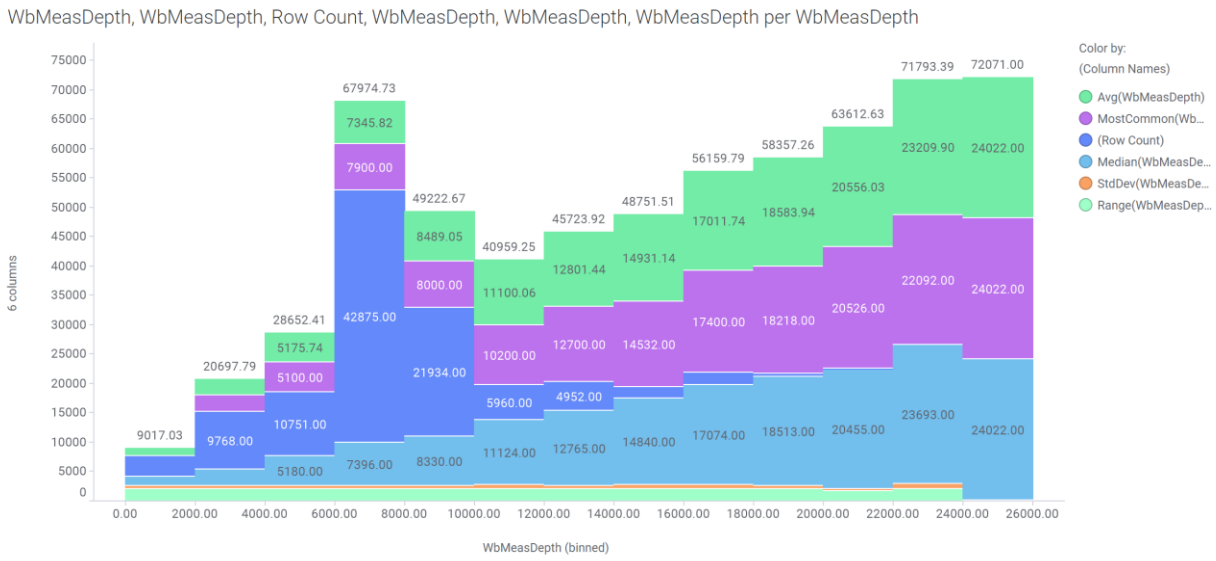




c. Plot number of wells against total depth (use Spotfire)



d. Run mean, median, mode, range and standard deviation on wells versus total depth (use website calculator)



e. Write one page paper on what you learned.

Ans- Tibco Spotfire is a fantastic tool that makes visual data science simpler, faster, and more auditable.

• This platform enables businesses to turn their data into meaningful insights fast and efficiently.

• Plotting dynamic data on maps is simple.

• We'll be able to develop actionable visual analytics right away.

• Spotfire makes organizations smarter, delivers AI-driven analytics, and makes it simple to plot interactive data on maps by allowing users to combine data in a single study and receive a holistic picture of the same with an interactive visualization.

The platform enables businesses to quickly and easily translate their data into actionable insights.

• The dashboard navigation and statistics are simple to use.

• It comes with cutting-edge collaborative tools.

• Data discovery and visualization are also possible.

• It has the ability to handle and analyze large amounts of data.

• It can quickly find outliers in large datasets.

• Compiling data and displaying it in a dynamic chart.

• New users will find it rather simple to get started

• Creating customizable interactive reports and dashboards.

• Handling many data sources is simple.

• When used in conjunction with the cloud version, users can access the library from anywhere with an internet connection.

Spotfire software makes it simple to plot dynamic data on maps and provides AI-driven insights to help businesses become more intelligent. Businesses may use the platform to rapidly and easily turn their data into actionable insights. It enables organizations to analyze data more quickly, confidently, and accurately.