**FIVE DATA VISUALIZATION CHARTS GENERATED FROM THE PLEIADAS**

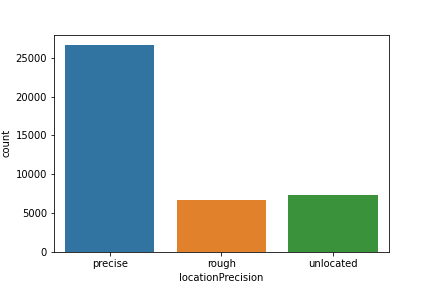
**DATASET.**

***1***

***Aim:*** To show the distribution of locationprecision in the dataset.

***Visual Design Type***: Count plot

***Visual Image:***

******

***Visual Mapping***: X axis contains the names of the locationprecision types(precise/rough) ;Y axis gives the count of the locationprecision types as it occurs in the dataset. Shape/size of image: 8\*8. Precise: representative point is an attested location. Rough: Representative point is the centroid of a bounding box. Unlocated: Representative point unknown

***Data preparation:*** The Data Cleaning and preparation performed to produce the above visualization was done using python pandas. The location dataset was downloaded from <http://atlantides.org/downloads/pleiades/dumps>. The dataset was loaded into pandas and cleaned by filling all the missing points in the locationprecision column as 'Unlocated', the matplotlib library was the used to create the count plot chart.

***Insight Generate:*** The chart shows the distribution of the locationprecisions in the dataset. The chart reveals that majority of the Locations in the dataset are precise while almost equal distribution occurs for rough and unlocated sites.

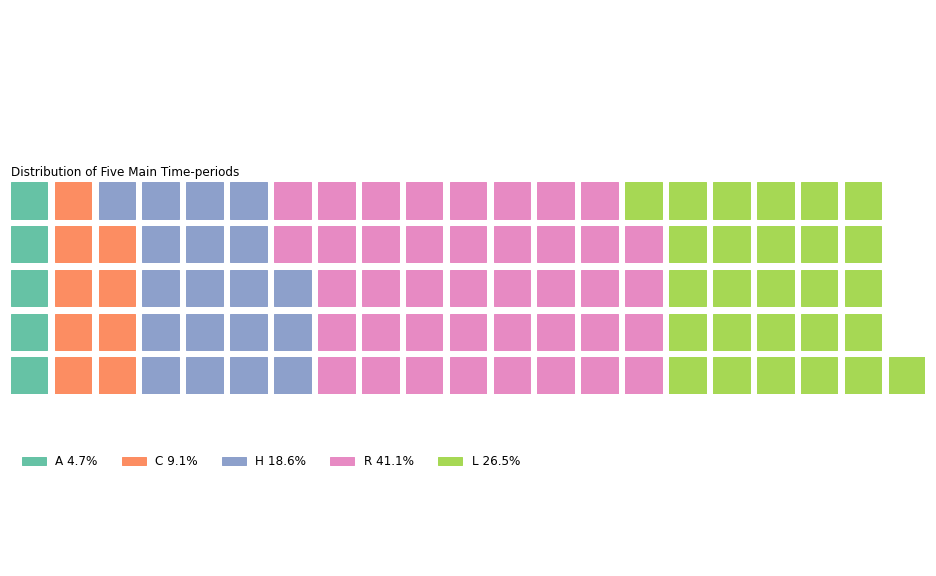
***2***

***Aim*:** To show the distribution of the different timeperiods of the Locations in the dataset.

***Visual Design Type***:Waffle Chart

***Visual Mapping***: Legends and Colour (A:Deep Green, H:Blue, L: Light Green, R:Purple, C: Orange). A' (1000-550 BC), 'C' (550-330 BC), 'H' (330-30 BC), 'R' (AD 30-300), 'L' (AD 300-640).

***Visual Image:***

******

***Data preparation***: The Data Cleaning and preparation performed to produce the above visualization was done using python pandas. The location dataset was downloaded from <http://atlantides.org/downloads/pleiades/dumps>. The dataset was loaded into pandas and a dictionary was created to take each of the respective timeperiods and number of times they occur in each location point. The timeperiods each location existed in was counted and updated Into the dictionary through a forward loop.

***Improvements***: The chat mainly shows the distribution of five main timeperiods, however several locations existed in other timeperiods apart from the five which could also have been included.

***Insight Generated:*** The chart shows the distribution of the five main timeperiods of locations in the dataset.The chart reveals that majority of the Locations existed in the R era followed by L; with very few existing till the A era.

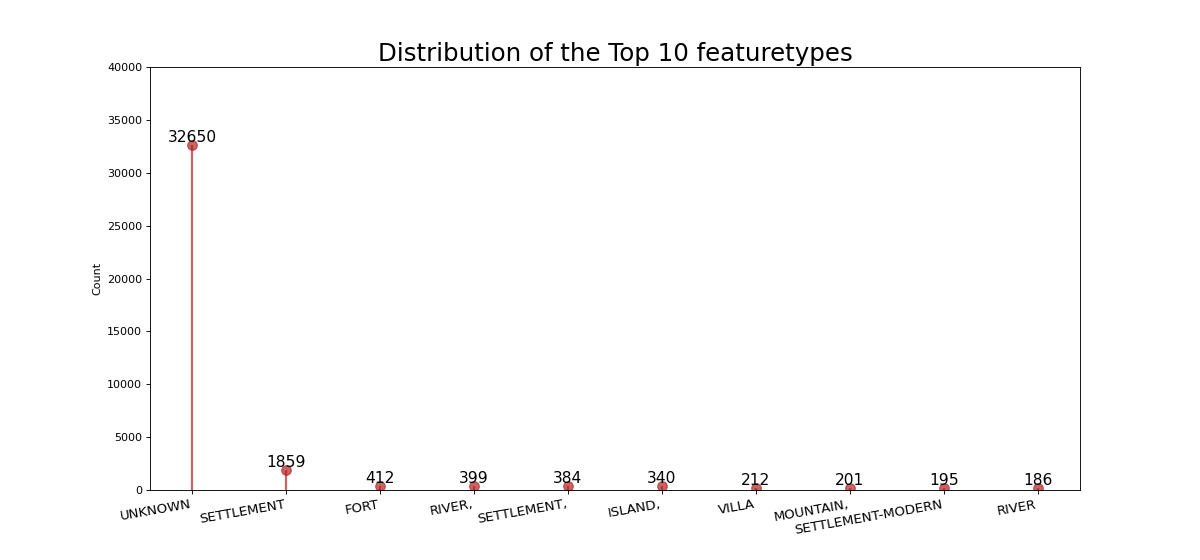
***3***

***Aim***: To show the distribution of 10 most occurring featuretypes of locations in the dataset.

***Visual Design Type***: Stem plot

***Visual Mapping:*** X axis contains the names of the featuretypes(settlement, temple,rivers etc)***,*** Y axis gives the count of the featuretypes as it occurs in the dataset.Shape/size of image: 12\*8.

***Visual Image***:



***Data preparation***: The Data Cleaning and preparation performed to produce the above visualization was done using python pandas. The location dataset was downloaded from http://atlantides.org/downloads/pleiades/dumps. The dataset was loaded into pandas and cleaned by filling all the missing values in the feature type column as 'Unknown'. The unique featuretypes and their respective number of occurrence was then extracted from the featuretype column and top 10 most occurring was plotted out using the matplotlib library.

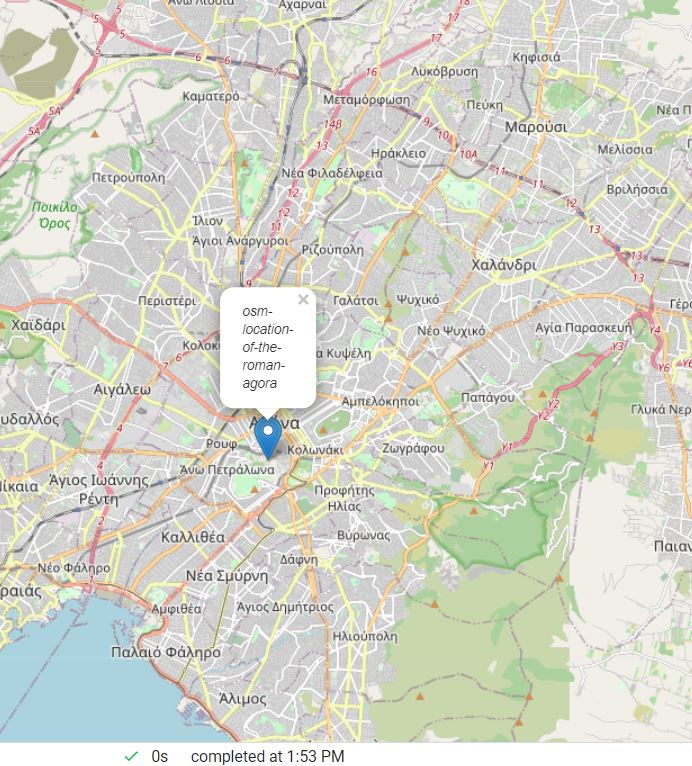
***Improvements:*** About 300 unique featuretypes was discovered and more than 10 could have been plotted out.Also more data cleaning and formatting could be done to put the dataset in appropriate shape for the Visualization as to avoid a unique value showing up with different names as seen in settlement/settlement,/settlement-modern.

***Insight Generated:*** The chart shows the top 10 most occurring featuretypes.From the chart we see that many location featuretypes are unknown and while about 1859 locations were settlements, 412 locations were forts and 340 were islands.

***4***

***Aim:*** To show the location of an extant historical site on the Map.

***Visual Image:***

******

***Data preparation:*** The Data Cleaning and preparation performed to produce the above visualization was done using python pandas.The location dataset was downloaded from <http://atlantides.org/downloads/pleiades/dumps>.The dataset was loaded into pandas and locations which posses "extant remains" as tags were extracted along with their respective longitude and latitude.Folium map is then used to create a map visualization of one of them (The Roman agora).

***Improvements:*** About 21 location posses’ extant remains, more could have been shown via same mapping.

***Insight Generated***: The present day location of the roman algora historical site.

***5***

***Aim:*** To show the Clustering relationship occuring amidst locations with extant remains based on their geographic coordinates.

***Visual Design Type***: Dendrogram (hierarchy clustering plot)

***Visual Mapping***: X axis contains the names of the Id of the locations, Y axis gives the Clustering degree that exists amidst this locations.

***Data preparation***: The Data Cleaning and preparation performed to produce the above visualization was done using python pandas. The location dataset was downloaded from http://atlantides.org/downloads/pleiades/dumps. The dataset was loaded into pandas and locations with "extant remains" as tag was extracted alongside their respective longitude and latitude. Using the longitude and latitude of the sites with extant remains a clustering was done and the result of the Clustering was visualized using a dendrogram.

***Improvements***: Proper Labelling of the X axis using the respective location id's rather than arbitrary numbers.

***Insight Generated:*** The chart clustering degree of extant sites based on geographic coordinates.From the chart we see how locations are close or far apart from each other. For example location 19 is farthest from 12, while 12 is closest to location 3.