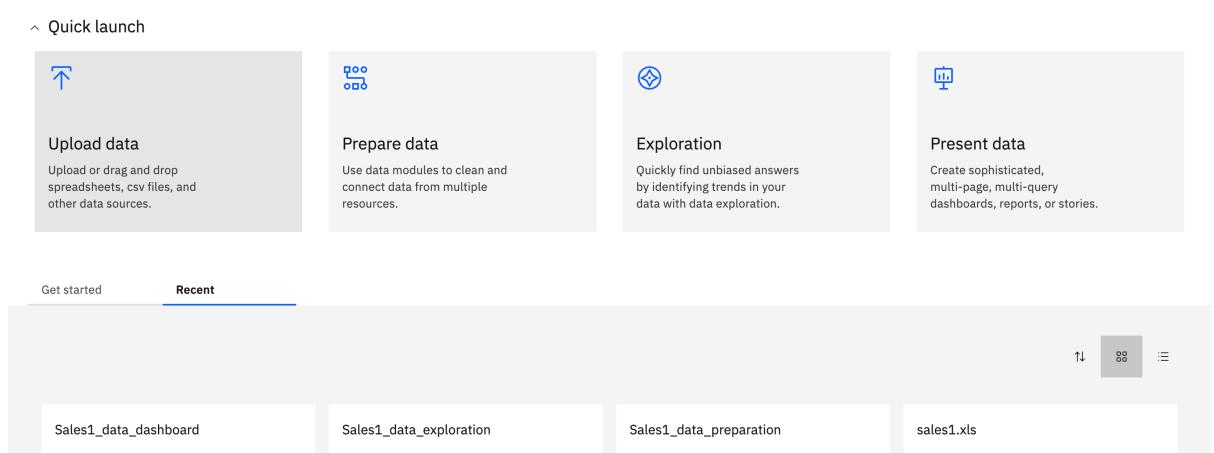
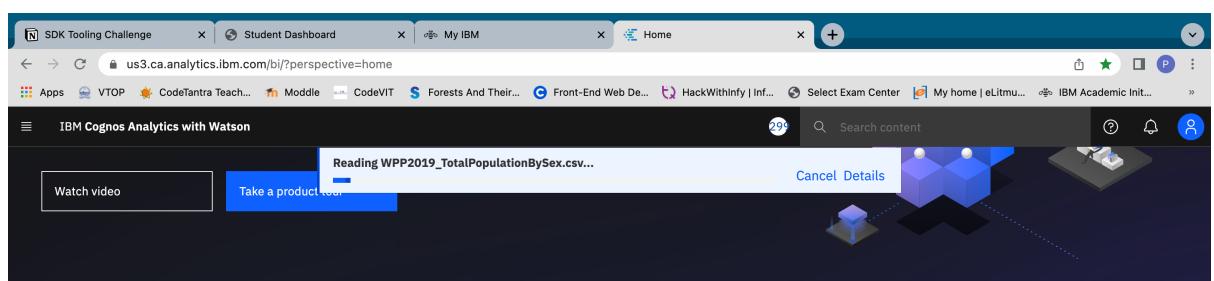
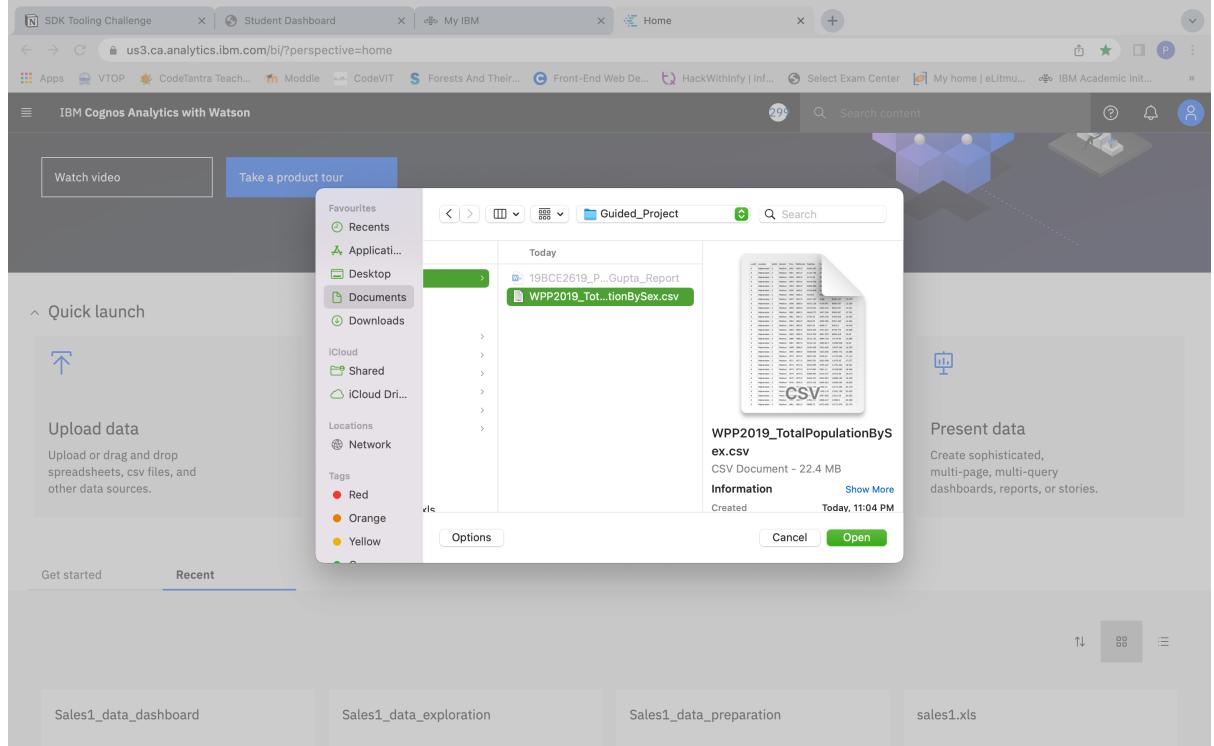


## Data Analytics Guided Project

### World Population Data Analytics Using IBM Cognos

#### 1. Loading the dataset



## 2. Preparing the dataset

The screenshot shows the IBM Cognos Analytics with Watson interface. On the left, there's a sidebar titled "Data module" with a search bar. Below it, a tree view shows a folder named "New data module" containing a file "WPP2019\_T...BySex.csv". This file has several columns listed under it: Row ID, LocID, Location, VarID, Variant, Time, MidPeriod, PopMale, PopFemale, PopTotal, and PopDensity. In the center, there's a "Grid" tab selected, showing a preview of the data. A circular icon with a grid and a checkmark is displayed. Below the grid, a section titled "Preview data" contains the text: "To preview data, select a table, a column in a table, or a folder that contains columns."

Fig: Dataset loaded

This screenshot shows the same interface as above, but with a context menu open over the "PopTotal" column in the grid. The menu options include "Filter...", "Create calculation...", "Create data group...", "Hide from users", "Remove", "Refresh properties...", "Format data...", "Clean...", "Rename", "Cut", "Copy", and "Properties". The "PopTotal" column header is highlighted with a blue border.

| Variant | Time | MidPeriod | PopMale  | PopFemale | PopTotal |
|---------|------|-----------|----------|-----------|----------|
| Medium  | 1950 | 1950.5    | 4099.243 | 3652.874  | 7752.117 |
| Medium  | 1951 | 1951.5    | 4134.756 | 3705.395  | 7840.151 |
| Medium  | 1952 | 1952.5    | 4174.45  | 3761.546  | 7935.996 |
| Medium  | 1953 | 1953.5    | 4218.336 | 3821.348  | 8039.684 |
| Medium  | 1954 | 1954.5    | 4266.484 | 3884.832  | 8151.316 |
| Medium  | 1955 | 1955.5    | 4318.945 | 3952.047  | 8270.992 |
| Medium  | 1956 | 1956.5    | 4375.8   | 4023.073  | 8398.873 |
| Medium  | 1957 | 1957.5    | 4437.157 | 4098      | 8535.157 |
| Medium  | 1958 | 1958.5    | 4503.156 | 4176.941  | 8680.097 |
| Medium  | 1959 | 1959.5    | 4573.914 | 4260.033  | 8833.947 |
| Medium  | 1960 | 1960.5    | 4649.573 | 4347.394  | 8996.967 |
| Medium  | 1961 | 1961.5    | 4730.25  | 4439.156  | 9169.406 |
| Medium  | 1962 | 1962.5    | 4816.05  | 4535.392  | 9351.442 |
| Medium  | 1963 | 1963.5    | 4907.03  | 4636.17   | 9543.2   |
| Medium  | 1964 | 1964.5    | 5003.245 | 4741.527  | 9744.772 |

Fig: Remove Population table

| Variant | Time | MidPeriod | PopMale  | PopFemale | PopDensity |
|---------|------|-----------|----------|-----------|------------|
| Medium  | 1950 | 1950.5    | 4099.243 | 3652.874  | 11.874     |
| Medium  | 1951 | 1951.5    | 4134.756 | 3705.395  | 12.009     |
| Medium  | 1952 | 1952.5    | 4174.45  | 3761.546  | 12.156     |
| Medium  | 1953 | 1953.5    | 4218.336 | 3821.348  | 12.315     |
| Medium  | 1954 | 1954.5    | 4266.484 | 3884.832  | 12.486     |
| Medium  | 1955 | 1955.5    | 4318.945 | 3952.047  | 12.669     |
| Medium  | 1956 | 1956.5    | 4375.8   | 4023.073  | 12.865     |
| Medium  | 1957 | 1957.5    | 4437.157 | 4098      | 13.073     |
| Medium  | 1958 | 1958.5    | 4503.156 | 4176.941  | 13.295     |
| Medium  | 1959 | 1959.5    | 4573.914 | 4260.033  | 13.531     |
| Medium  | 1960 | 1960.5    | 4649.573 | 4347.394  | 13.781     |
| Medium  | 1961 | 1961.5    | 4730.25  | 4439.156  | 14.045     |
| Medium  | 1962 | 1962.5    | 4816.05  | 4535.392  | 14.324     |
| Medium  | 1963 | 1963.5    | 4907.03  | 4636.17   | 14.618     |
| Medium  | 1964 | 1964.5    | 5003.245 | 4741.527  | 14.926     |

Fig: Dataset after removing TotalPopulation Column

Create calculation

Name: PopTotal

Components:

- WPP2019\_To...BySex.csv
- # Row Id
- # LocID
- #c Location
- # VarID
- #c Variant
- Time
- MidPeriod
- PopMale
- PopFemale
- PopDensity

Expression:

```
1 PopMale + PopFemale
```

Validation Results:

The expression is valid.

Cancel OK

Fig: Creating PopTable combining PopMale and PopFemale

|  | PopTotal          | Row Id | LocID | Location    | VarID | Variant | Tin  |
|--|-------------------|--------|-------|-------------|-------|---------|------|
|  | 7752.117          | 1      | 4     | Afghanistan | 2     | Medium  | 1950 |
|  | 7840.151          | 2      | 4     | Afghanistan | 2     | Medium  | 1951 |
|  | 7935.995999999999 | 3      | 4     | Afghanistan | 2     | Medium  | 1952 |
|  | 8039.684          | 4      | 4     | Afghanistan | 2     | Medium  | 1953 |
|  | 8151.31600000001  | 5      | 4     | Afghanistan | 2     | Medium  | 1954 |
|  | 8270.992          | 6      | 4     | Afghanistan | 2     | Medium  | 1955 |
|  | 8398.873          | 7      | 4     | Afghanistan | 2     | Medium  | 1956 |
|  | 8535.157          | 8      | 4     | Afghanistan | 2     | Medium  | 1957 |
|  | 8680.097          | 9      | 4     | Afghanistan | 2     | Medium  | 1958 |
|  | 8833.947          | 10     | 4     | Afghanistan | 2     | Medium  | 1959 |
|  | 8996.967          | 11     | 4     | Afghanistan | 2     | Medium  | 1960 |
|  | 9169.405999999999 | 12     | 4     | Afghanistan | 2     | Medium  | 1961 |
|  | 9351.442          | 13     | 4     | Afghanistan | 2     | Medium  | 1962 |
|  | 9543.2            | 14     | 4     | Afghanistan | 2     | Medium  | 1963 |
|  | 9744.772          | 15     | 4     | Afghanistan | 2     | Medium  | 1964 |

Fig: PopTotal Created

| Location    | VarID | Variant | Time | MidPeriod | PopMale  |
|-------------|-------|---------|------|-----------|----------|
| Afghanistan | 2     | Medium  | 1950 | 1950.5    | 4099.243 |
| Afghanistan | 2     | Medium  | 1951 | 1951.5    | 4134.756 |
| Afghanistan | 2     | Medium  | 1952 | 1952.5    | 4174.45  |
|             | 2     | Medium  | 1953 | 1953.5    | 4218.336 |
|             | 2     | Medium  | 1954 | 1954.5    | 4266.484 |
|             | 2     | Medium  | 1955 | 1955.5    | 4318.945 |
|             | 2     | Medium  | 1956 | 1956.5    | 4375.8   |
|             | 2     | Medium  | 1957 | 1957.5    | 4437.157 |
|             | 2     | Medium  | 1958 | 1958.5    | 4503.156 |
|             | 2     | Medium  | 1959 | 1959.5    | 4573.914 |
|             | 2     | Medium  | 1960 | 1960.5    | 4649.573 |
|             | 2     | Medium  | 1961 | 1961.5    | 4730.25  |
|             | 2     | Medium  | 1962 | 1962.5    | 4816.05  |
|             | 2     | Medium  | 1963 | 1963.5    | 4907.03  |
|             | 2     | Medium  | 1964 | 1964.5    | 5003.245 |

Fig: Formatting PopMale column

**Data format**

**Column:** PopMale

**Format type:** Number

| Period | PopMale  |
|--------|----------|
| 0.5    | 4099.243 |
| 1.5    | 4134.756 |
| 2.5    | 4174.45  |
| 3.5    | 4218.336 |
| 4.5    | 4266.484 |
| 5.5    | 4318.945 |
| 6.5    | 4375.8   |
| 7.5    | 4437.157 |
| 8.5    | 4503.156 |
| 9.5    | 4573.914 |
| 0.5    | 4649.573 |
| 1.5    | 4730.25  |
| 2.5    | 4816.05  |
| 3.5    | 4907.03  |
|        | 5003.245 |

**Advanced options**

**Cancel** **OK**

Fig: Formatting PopMale column (Choosing Format Type)

**Grid** **Relationships** **Custom tables**

| VarID | Variant | Time | MidPeriod | PopMale   | PopFemale |
|-------|---------|------|-----------|-----------|-----------|
| 2     | Medium  | 1950 | 1950.5    | 4,099.243 | 3652.874  |
| 2     | Medium  | 1951 | 1951.5    | 4,134.756 | 3705.395  |
| 2     | Medium  | 1952 | 1952.5    | 4,174.45  | 3761.546  |
| 2     | Medium  | 1953 | 1953.5    | 4,218.336 | 3821.348  |
| 2     | Medium  | 1954 | 1954.5    | 4,266.484 | 3884.832  |
| 2     | Medium  | 1955 | 1955.5    | 4,318.945 | 3952.047  |
| 2     | Medium  | 1956 | 1956.5    | 4,375.8   | 4023.073  |
| 2     | Medium  | 1957 | 1957.5    | 4,437.157 | 4098      |
| 2     | Medium  | 1958 | 1958.5    | 4,503.156 | 4176.941  |
| 2     | Medium  | 1959 | 1959.5    | 4,573.914 | 4260.033  |
| 2     | Medium  | 1960 | 1960.5    | 4,649.573 | 4347.394  |
| 2     | Medium  | 1961 | 1961.5    | 4,730.25  | 4439.156  |
| 2     | Medium  | 1962 | 1962.5    | 4,816.05  | 4535.392  |
| 2     | Medium  | 1963 | 1963.5    | 4,907.03  | 4636.17   |
| 2     | Medium  | 1964 | 1964.5    | 5,003.245 | 4741.527  |

**Filter...**

**Format data...**

**Clean...**

**Rename...**

**Cut...**

**Copy...**

**Properties...**

Fig: Formatting PopFemale column

The screenshot shows the 'Data format' dialog for the 'PopFemale' column. The 'Format type:' dropdown is open, showing options like Unformatted, Text, Number, Percent, Currency, Date, Date/Time, Time, Time interval, and Custom. The 'Number' option is selected and highlighted with a green background. To the right of the dialog, a preview table shows data for the 'Male' and 'PopFemale' columns.

| Male      | PopFemale |
|-----------|-----------|
| 99,243    | 3652.874  |
| 84,756    | 3705.395  |
| 74,45     | 3761.546  |
| 18,336    | 3821.348  |
| 56,484    | 3884.832  |
| 18,945    | 3952.047  |
| 75,8      | 4023.073  |
| 37,157    | 4098      |
| 33,156    | 4176.941  |
| 73,914    | 4260.033  |
| 19,573    | 4347.394  |
| 30,25     | 4439.156  |
| 16,05     | 4535.392  |
| 37,03     | 4636.17   |
| 5,003,245 | 4741.527  |

Fig: Formatting PopFemale column (Chosing Format Type)

The screenshot shows the 'Select sources' dialog. On the left, there's a list of files under 'My content'. A red oval highlights the file 'WPP2019\_TotalPopulationBySex\_Prepared'. On the right, there are filters for 'Type' (Folders, Packages, Files, Data sets, Modules) and 'Modified' (All, Today, Yesterday, Past week, Past month). The 'All' option is selected under 'Modified'.

Fig: Prepared Data Module Saved

### 3. Data exploration

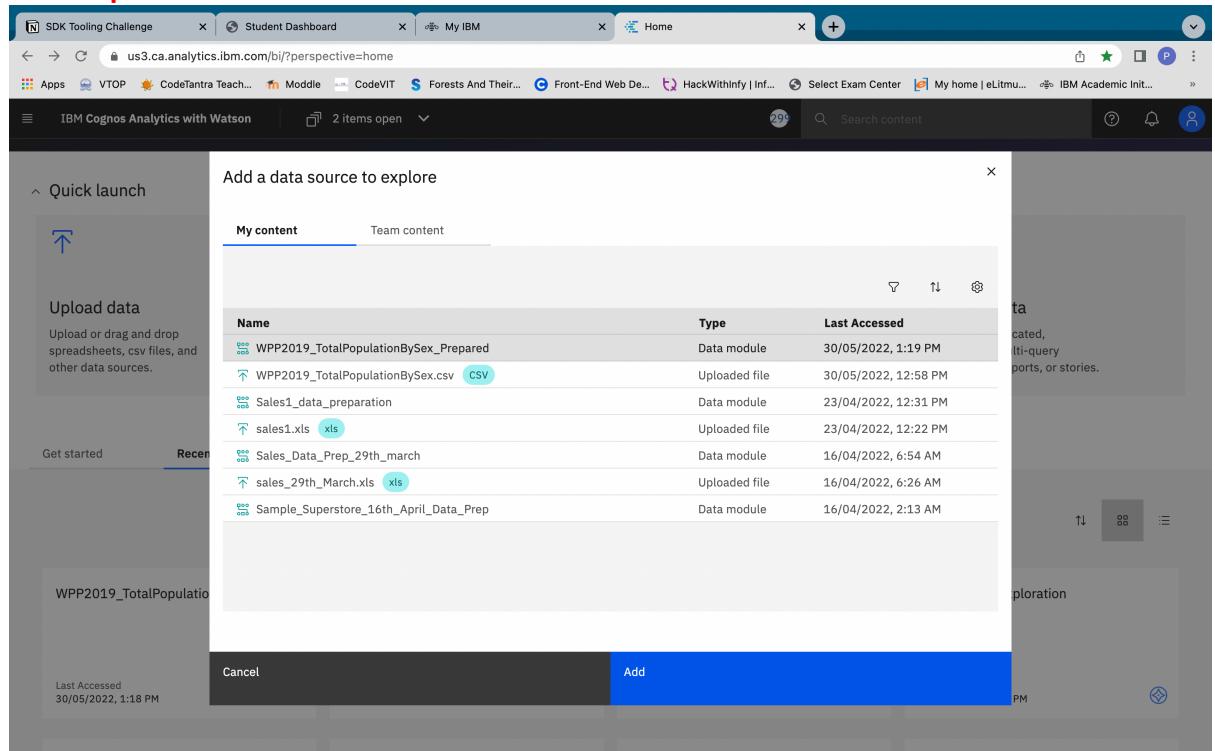


Fig: Loading data for data exploration

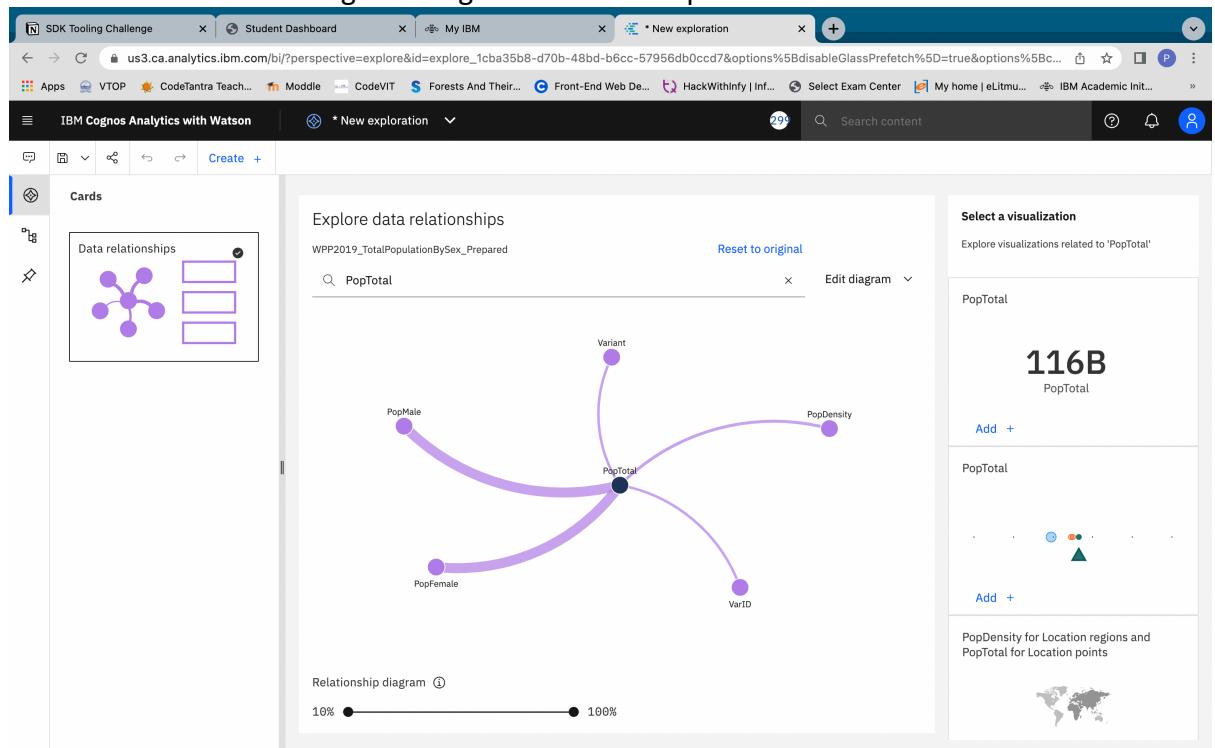


Fig: Loaded data for data exploration

The screenshot shows the IBM Cognos Analytics with Watson interface. On the left, there's a sidebar with 'Cards' and 'Data relationships'. In the center, a 'Relationship diagram' is displayed with nodes like 'PopTotal', 'PopFemale', 'PopDensity', 'VarID', and 'Variant'. A 'Relationship diagram' slider at the bottom ranges from 10% to 100%. To the right, there are sections for 'Select a visualization' showing 'PopTotal' with a value of '116B' and a world map icon.

Fig: Selecting single visualization

The screenshot shows the 'Choose visualization type' dialog in IBM Cognos Analytics with Watson. It lists various visualization types: Comparison (Bar, Bullet, Column, Heat map, Line and column, Marimekko, Radar, Stacked bar, Stacked column, Word cloud), Parts to whole (Hierarchy bubble, Packed bubble, Pie, Tree map), Trend (Area, Box plot, Line, Line and column, Point, Radial, Waterfall), Relationships (Bubble, Heat map, Network, Scatter, Crosstab, List, Summary, Table), Tables and summary (Geospatial, Other), Advanced analytics, and Geospatial.

Fig: Selecting Tree Map

19BCE2619  
Pranjal Gupta

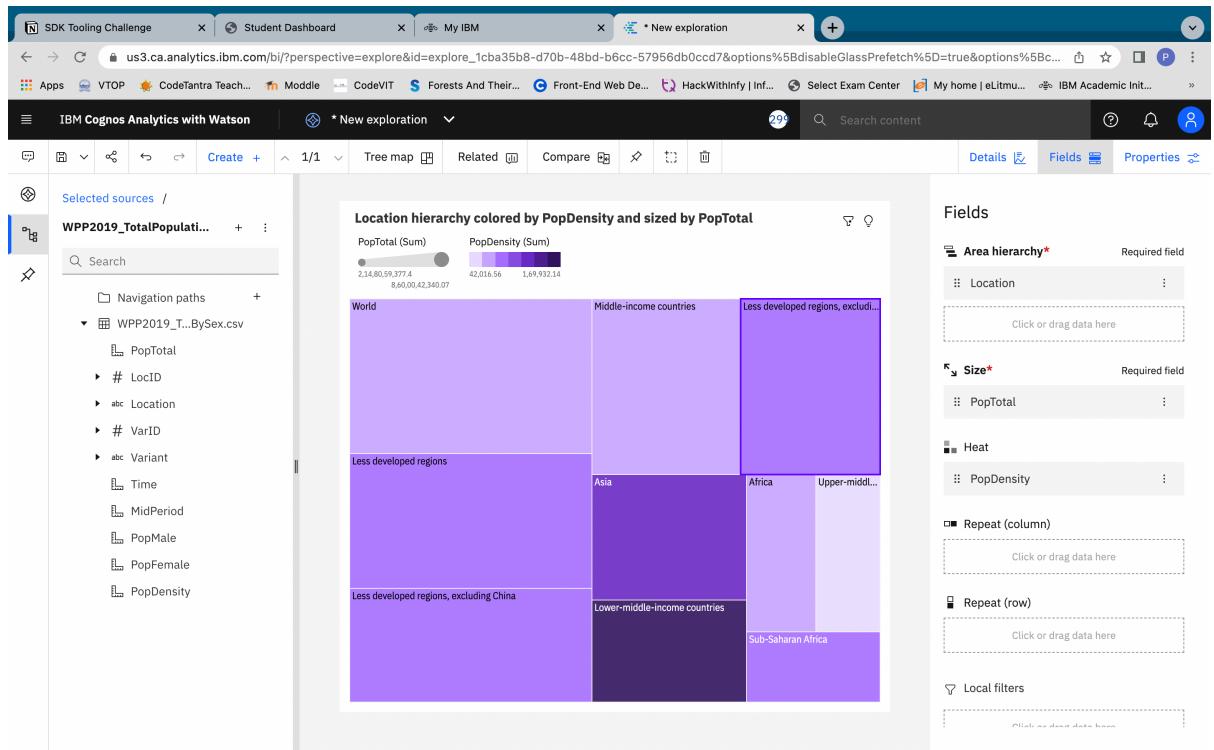


Fig: Top10 PopTotal By Location Using Tree Map

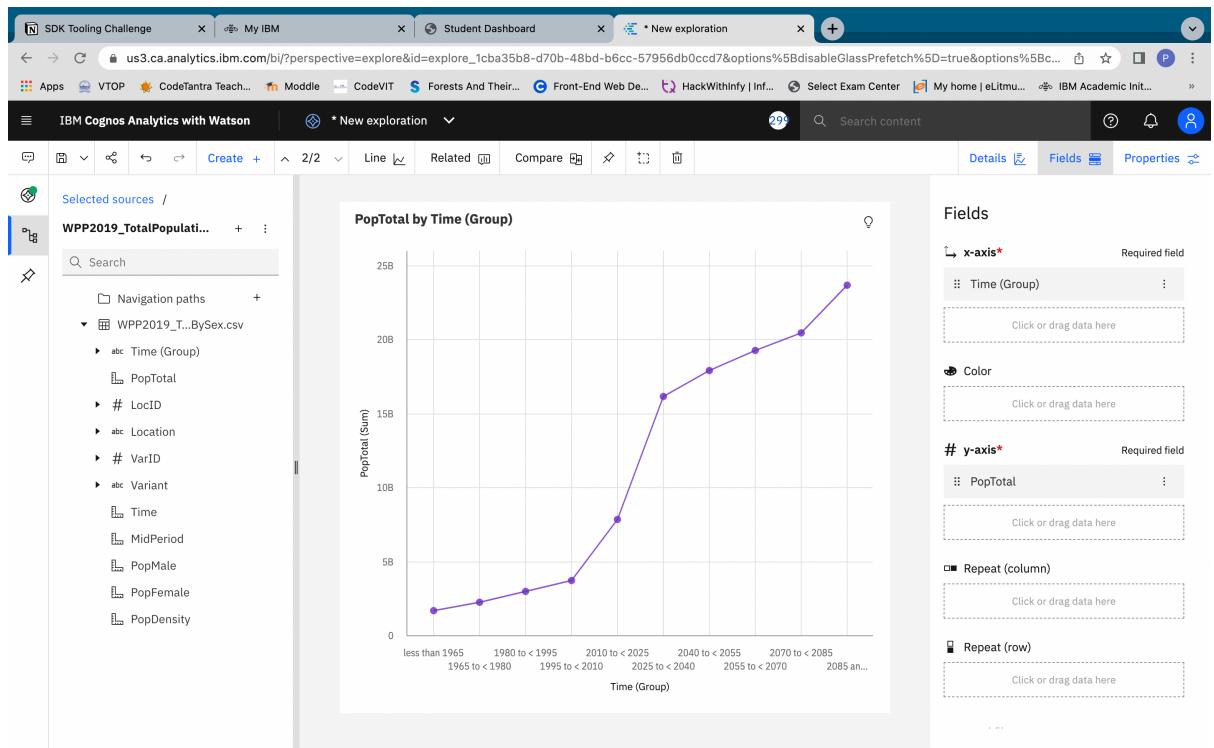


Fig: PopTotal by Time using Line Chart

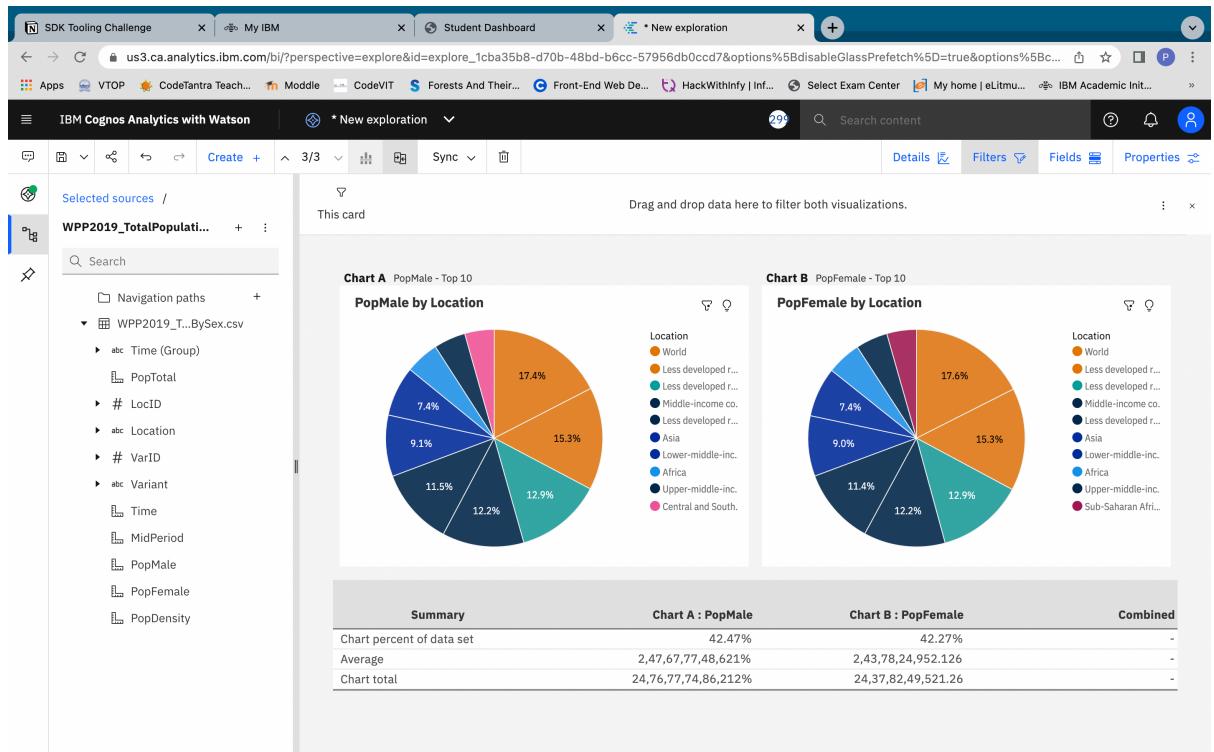


Fig: Pop Male by Location and Pop Female by Location Using Pie Charts

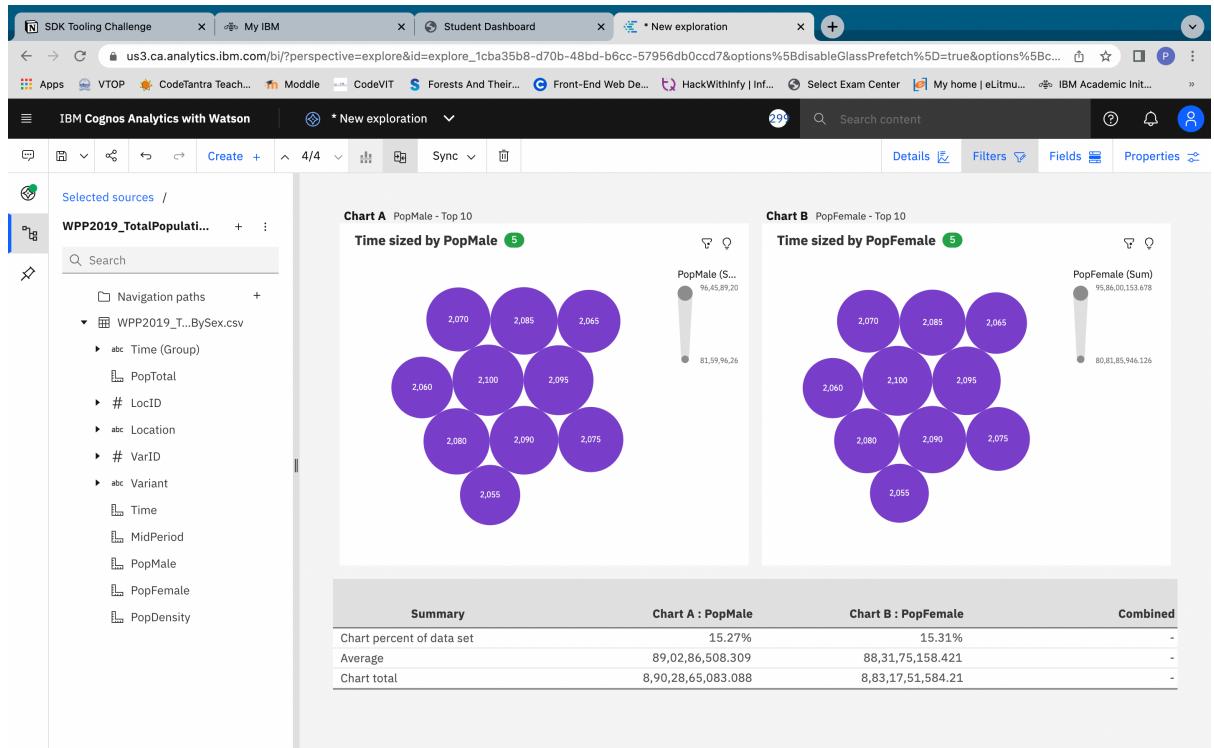


Fig: Pop Male By Time And Pop Female Using Packed Bubble Charts

#### 4. Preparing Dashboards

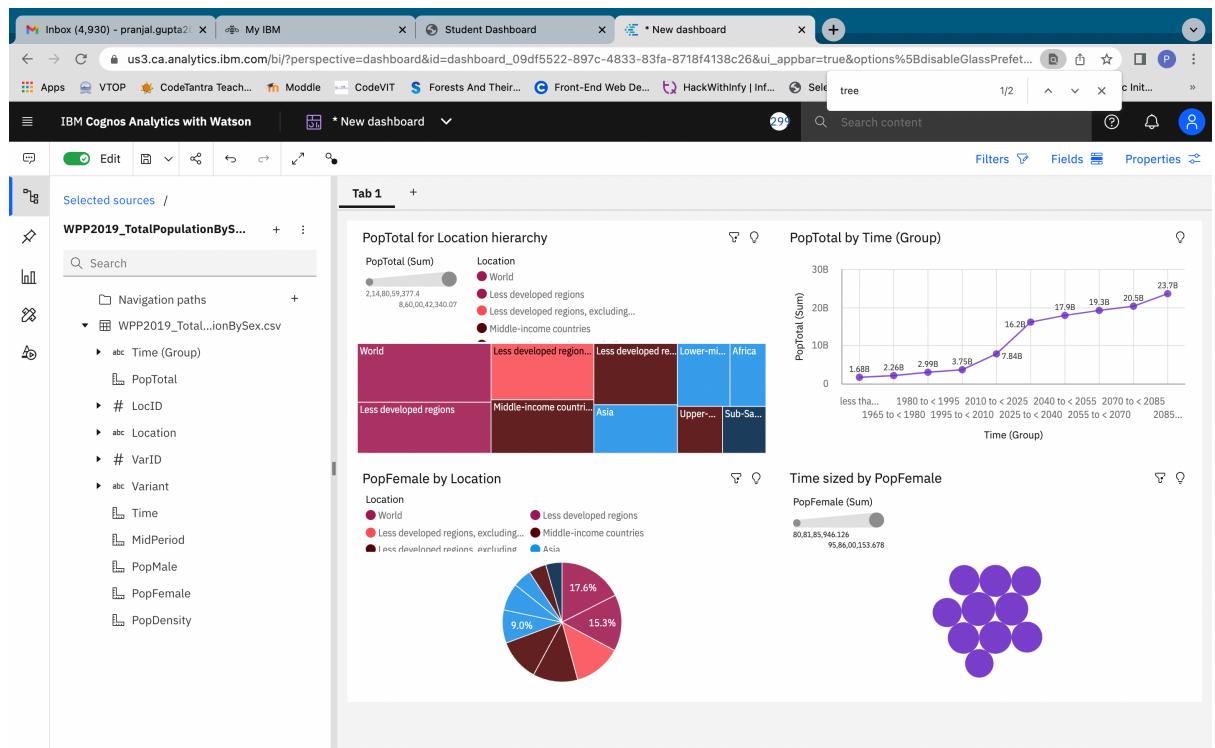


Fig: Dashboard 1

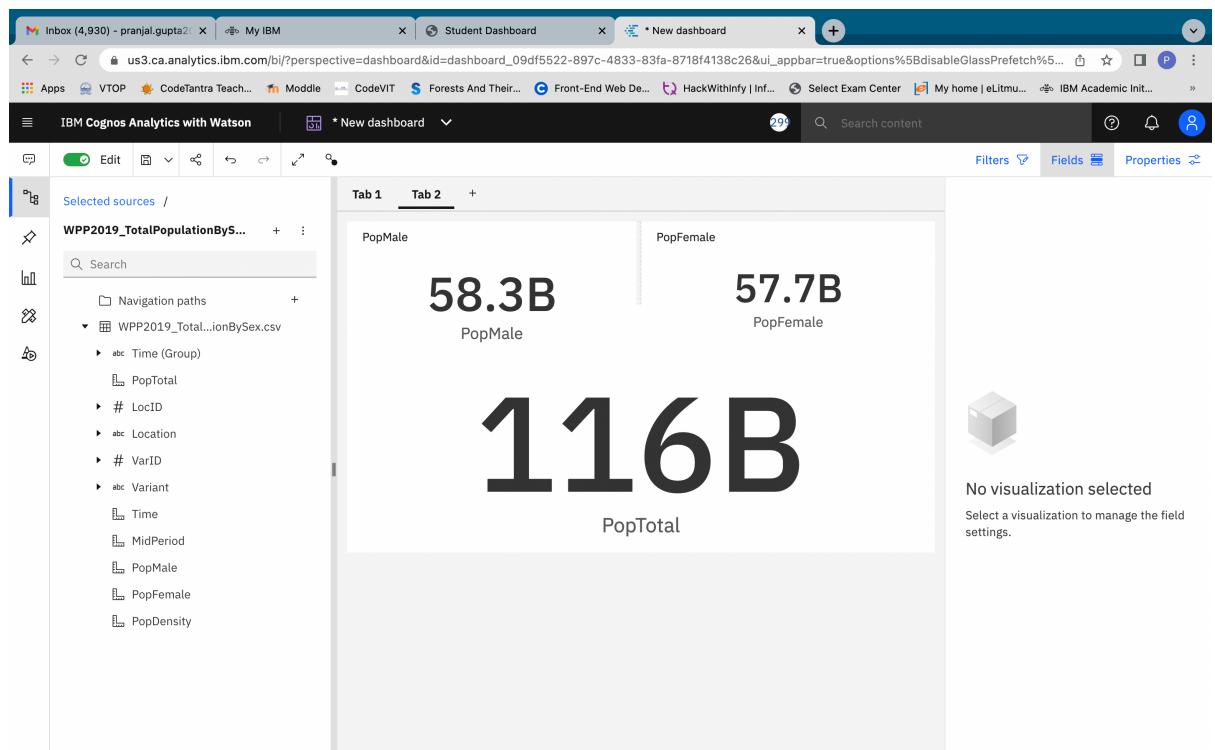


Fig: Dashboard 2: PopMale, PopFemale and PopTotal