PHASE 3:

**TN Marginal Workers Assessment**

**1. Load the Dataset:**

First, download the dataset from the provided link: <https://tn.data.gov.in/resource/marginal-workers-classified-age-industrial-category-and-sex-scheduled-caste-2011-tamil>.

The code for load a csv dataset files for analysing

import pandas as pd

data = pd.read\_csv("marginal\_workers\_tamil.csv")

**2. Data Pre-processing:**

Explore the dataset to understand its structure, and handle any missing or irrelevant data.

Python code for show on some data’s

print(data.head())

data.dropna(inplace=True)

selected\_columns = data[['age', 'industrial\_category']]

**3. Data Transformation:**

- You may need to convert categorical variables into numerical format for clustering.

- In this case, you might use one-hot encoding to convert the "industrial\_category" column into numerical features.

Code for encoding column :

data\_encoded = pd.get\_dummies(selected\_columns, columns=['industrial\_category'])

**4. Data Visualization :**

A bar chart, a pie chart, and heatmap visualization are common types of data visualization techniques used in data analysis and presentation.

**Bar Chart**:

A bar chart is a graphical representation of data in which rectangular bars of varying heights or lengths are used to represent different categories or values. Bar charts are commonly used to compare data between different categories.

**Pie Chart** **:**

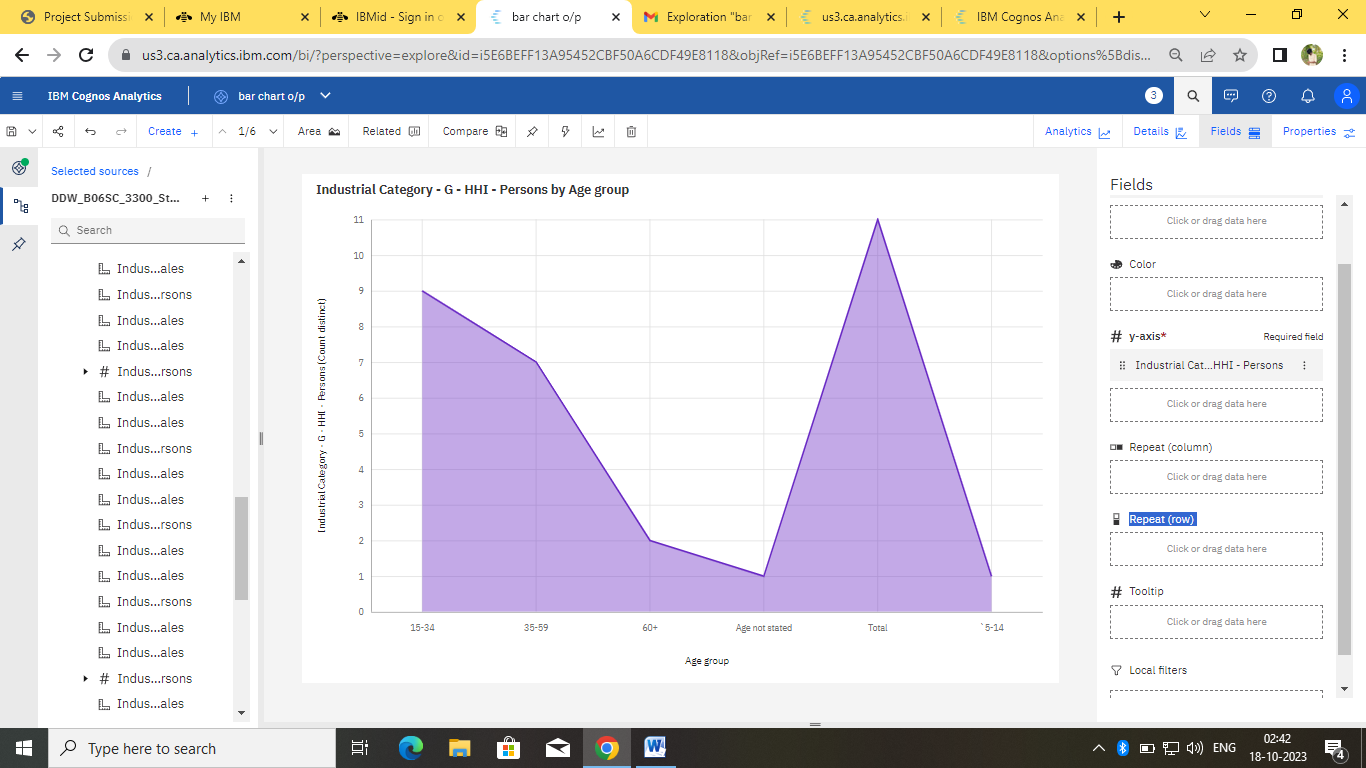
A pie chart is a circular chart divided into slices, where each slice represents a proportion of a whole. It's often used to show the composition or distribution of a single dataset in relation to the whole.

**Headmap:**

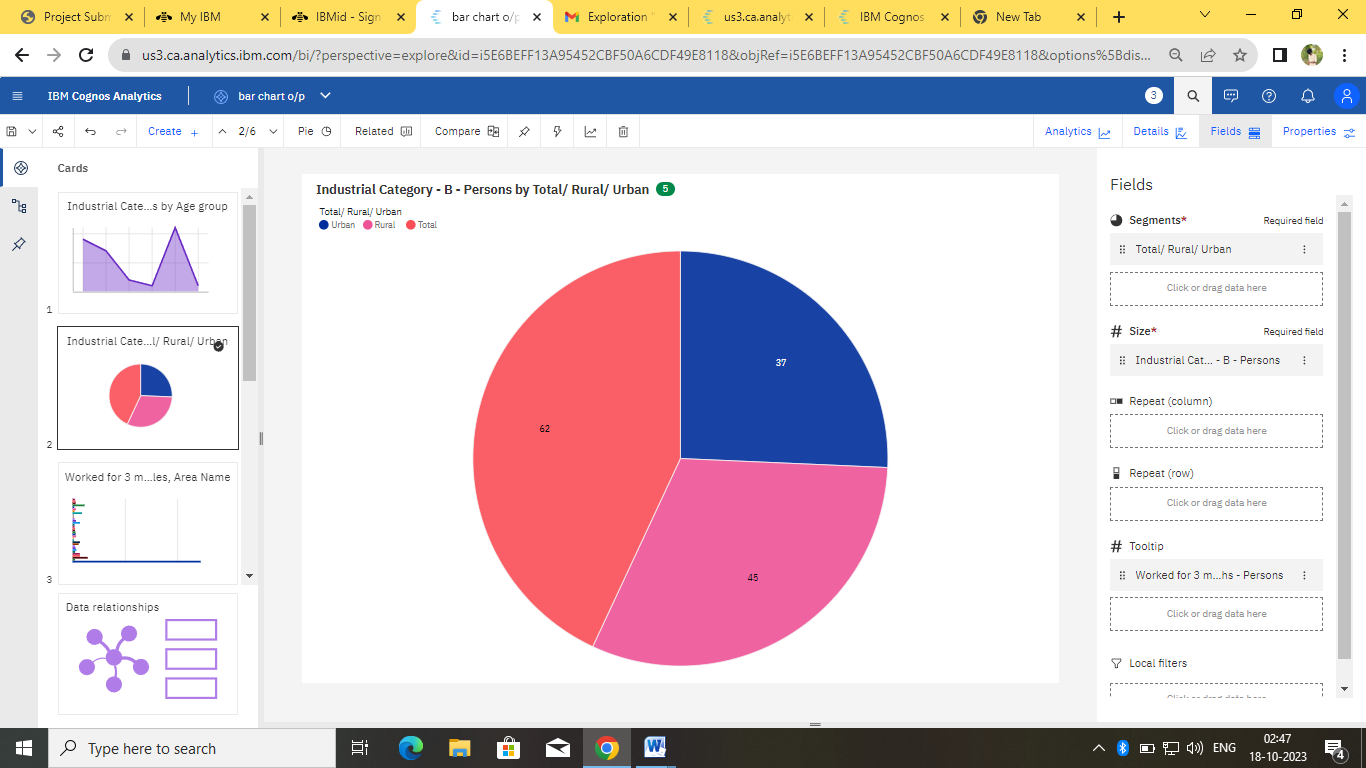
A heatmap is a two-dimensional representation of data where values in a matrix are represented using colors. It's commonly used to visualize relationships or patterns in data, such as correlation matrices.

**Data Visualization:**

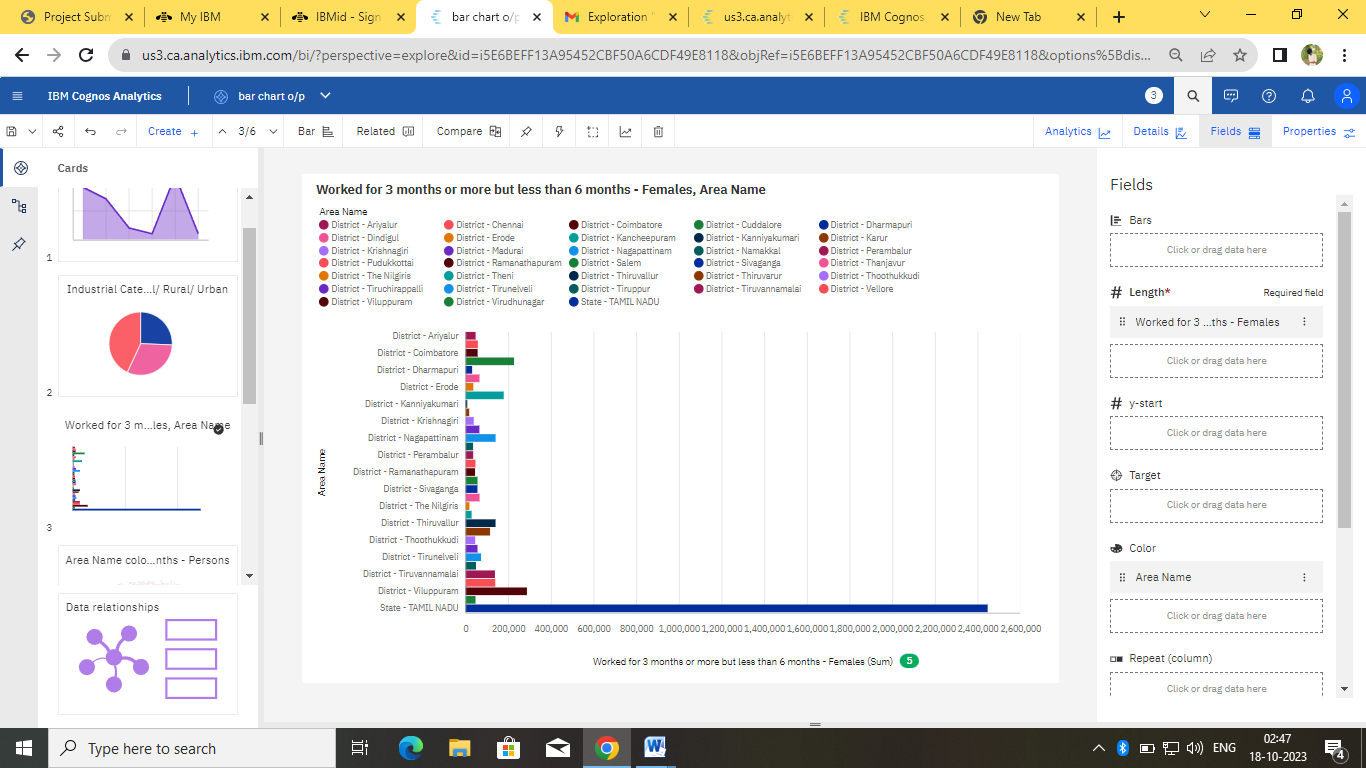
**Area graph**



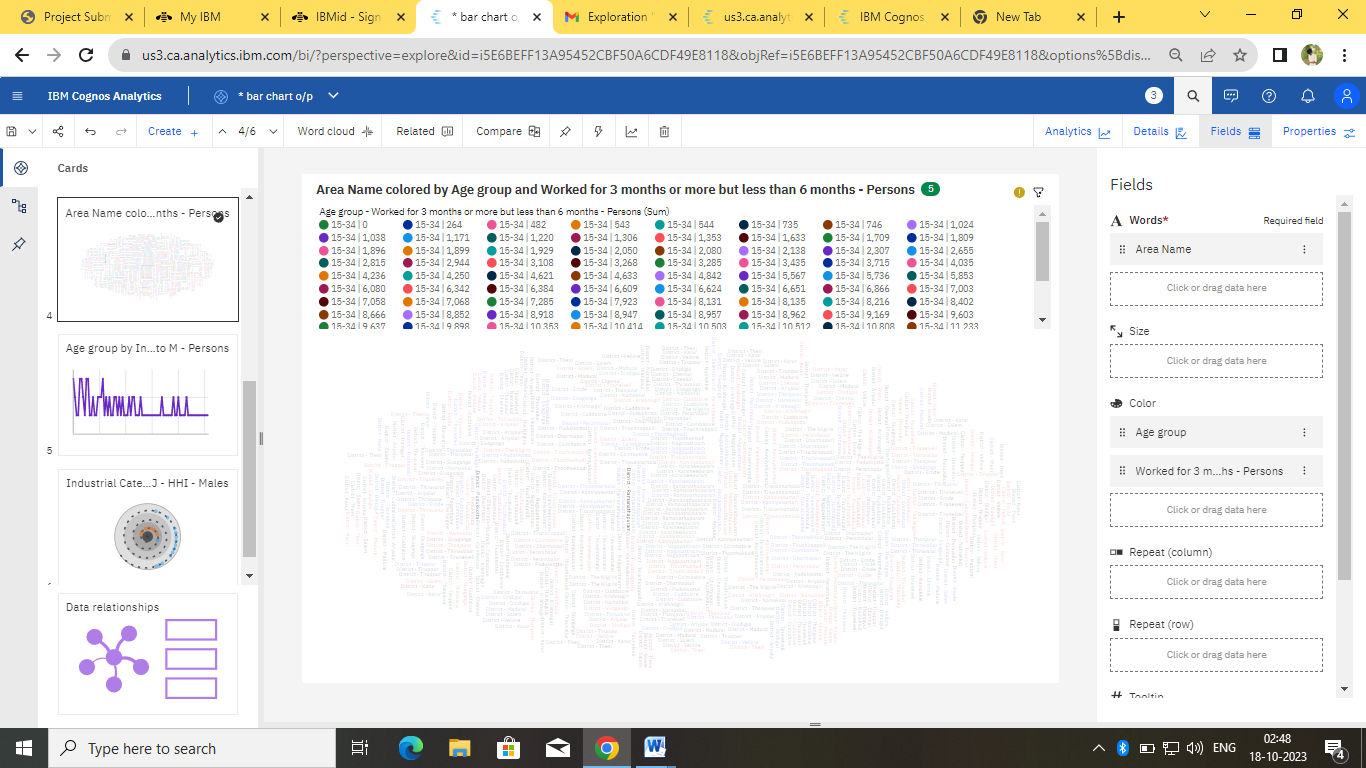
Pie chart:



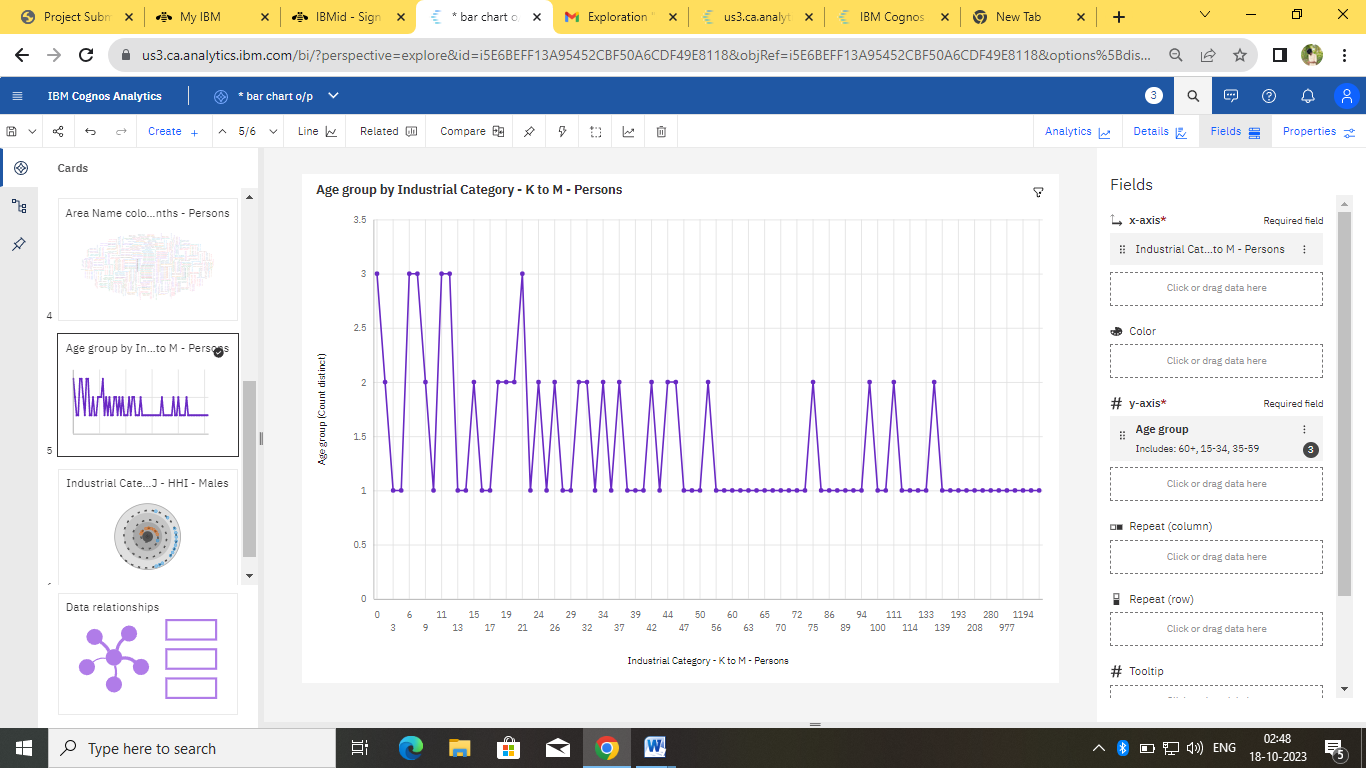
Bar chart:



**World cloud:**



Line chart:



Sprial :

