PUBLIC HEALTH CARE AWARENESS-DATASET IN VISUALIZATION

Step 1: Dataset Download

Access a healthcare dataset containing relevant information about healthcare-related metrics, such as patient information, medical procedures, diagnoses, outcomes, and other healthcare-specific data. You can obtain such datasets from sources like healthcare organizations, government health agencies, or research institutions.

Step 2: Loading the Dataset

Import the necessary Python libraries, including Pandas, for data analysis.

Load the healthcare dataset into IBM Cognos Analytics for further analysis.

Display the first few rows of the dataset to inspect the data structure.

Step 3: Exploratory Data Analysis

Perform an initial exploration of the healthcare dataset to understand its structure and features.

Check for missing values, data types, and basic statistics.

Visualize key features to gain insights into the data, such as patient demographics, medical procedure trends, or disease prevalence.

Identify potential relationships or correlations between features and relevant healthcare outcomes.

Step 4: Define Analysis Objectives

Define the objectives of the analysis for this phase, which may include:

Understanding patient demographics and their impact on healthcare outcomes.

Identifying key factors influencing medical procedure utilization.

Analyzing disease prevalence and trends.

Detecting anomalies or unusual healthcare patterns.

Step 5: Data Cleaning and Preprocessing

Clean and preprocess the healthcare data to ensure its quality and suitability for analysis:

Handle missing values by imputing or removing them based on data exploration.

Encode categorical variables (e.g., diagnosis codes) if necessary.

Perform feature engineering (create new features) if required.

If you want to predict certain healthcare outcomes, create a target variable, if not already present.

Step 6: IBM Cognos for Visualization

Utilize IBM Cognos for creating data visualizations that provide insights into the healthcare dataset:

Create various visualizations, such as bar charts, line charts, heatmaps, and interactive dashboards.

Visualize patient demographics, medical procedure trends, disease prevalence, or other relevant healthcare metrics.

Explore relationships between variables using scatter plots, correlation matrices, or other appropriate visualizations.

Step 7: Data Validation

Validate the processed data to ensure its quality and accuracy:

Perform data validation checks to confirm data consistency and accuracy.

Identify and address any potential data quality issues, such as outliers or data entry errors.

Step 8: Documentation

Maintain comprehensive documentation of the activities performed in this phase:

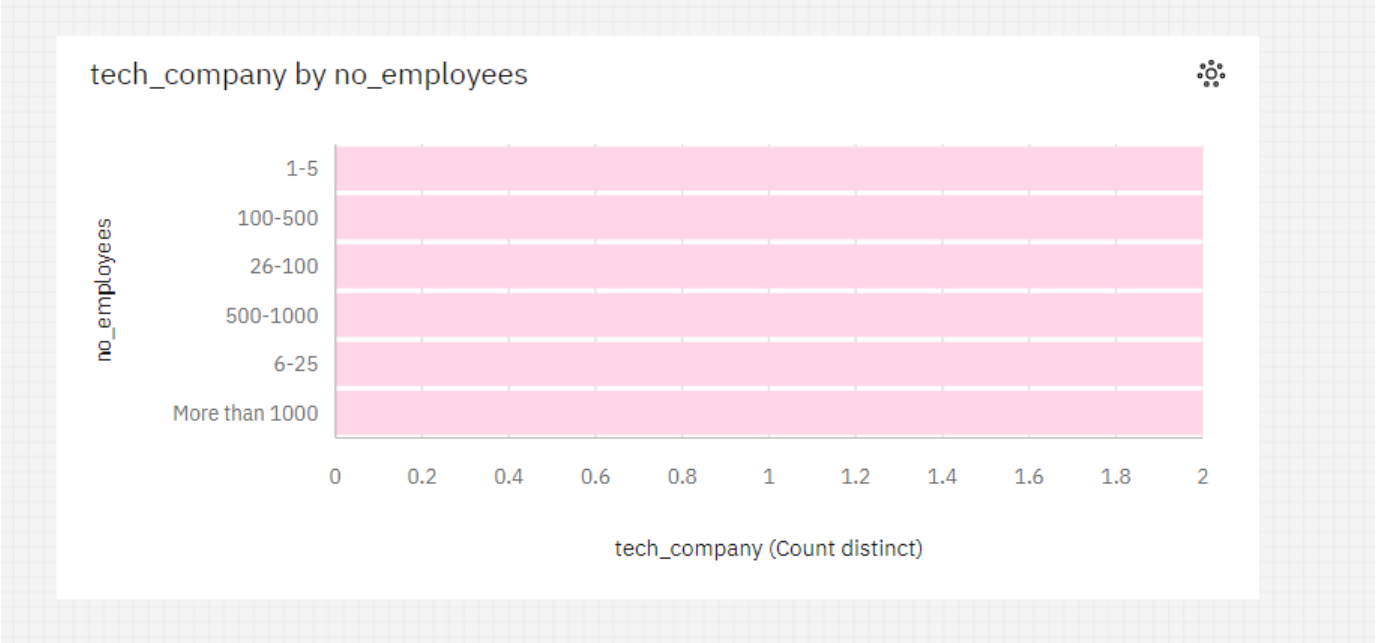
Record the data preprocessing steps, including any transformations and cleaning.

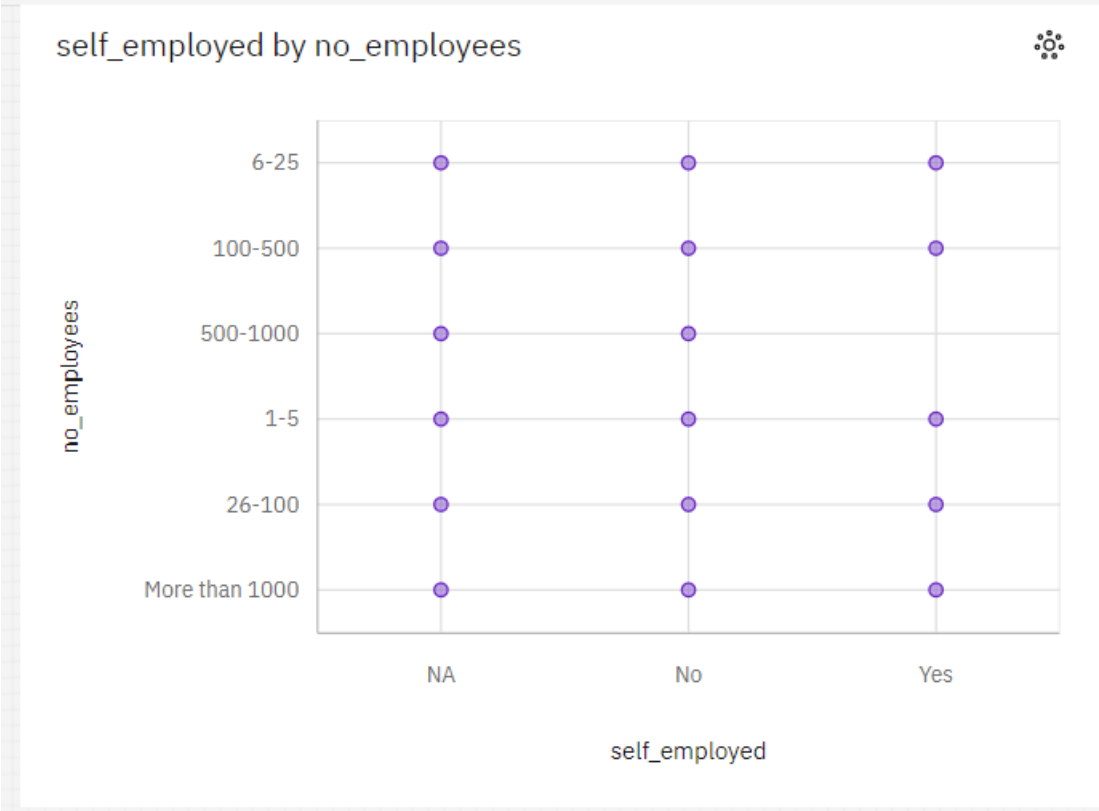
Document the objectives defined for this phase.

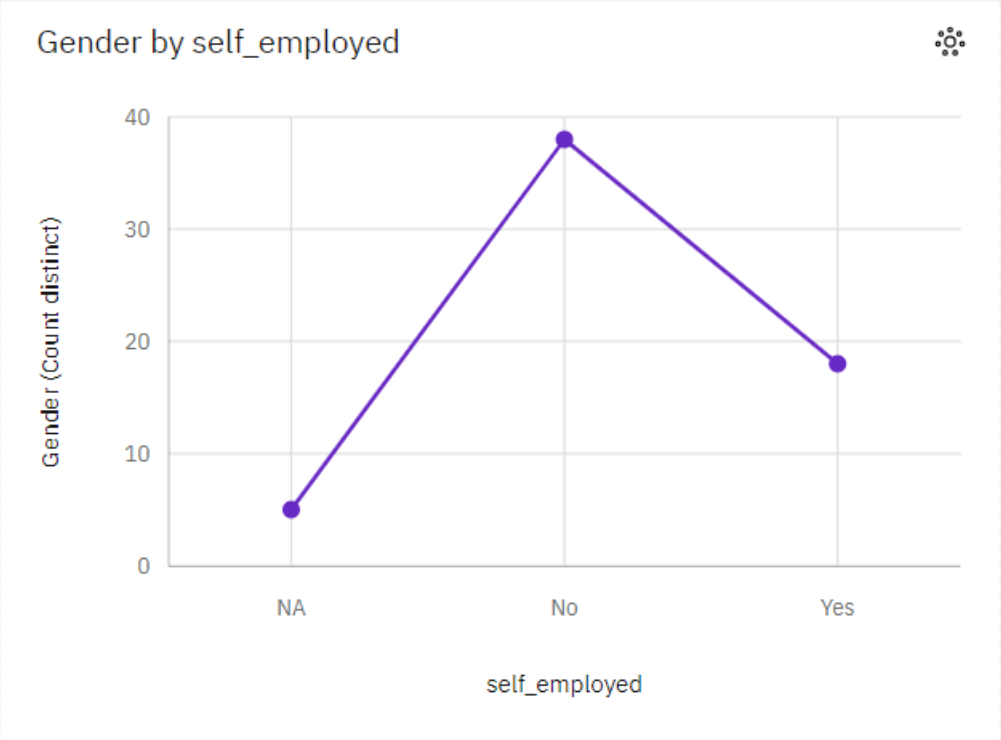
Capture insights gained from data visualization using IBM Cognos.

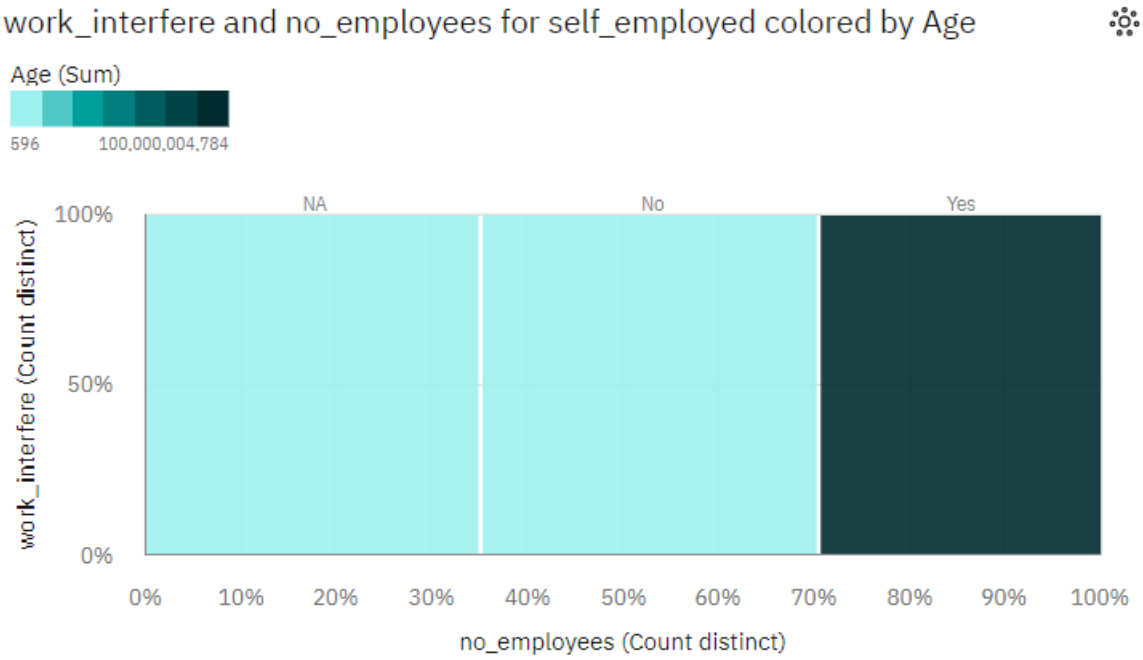
Document any observations, challenges, or discoveries made during this phase.

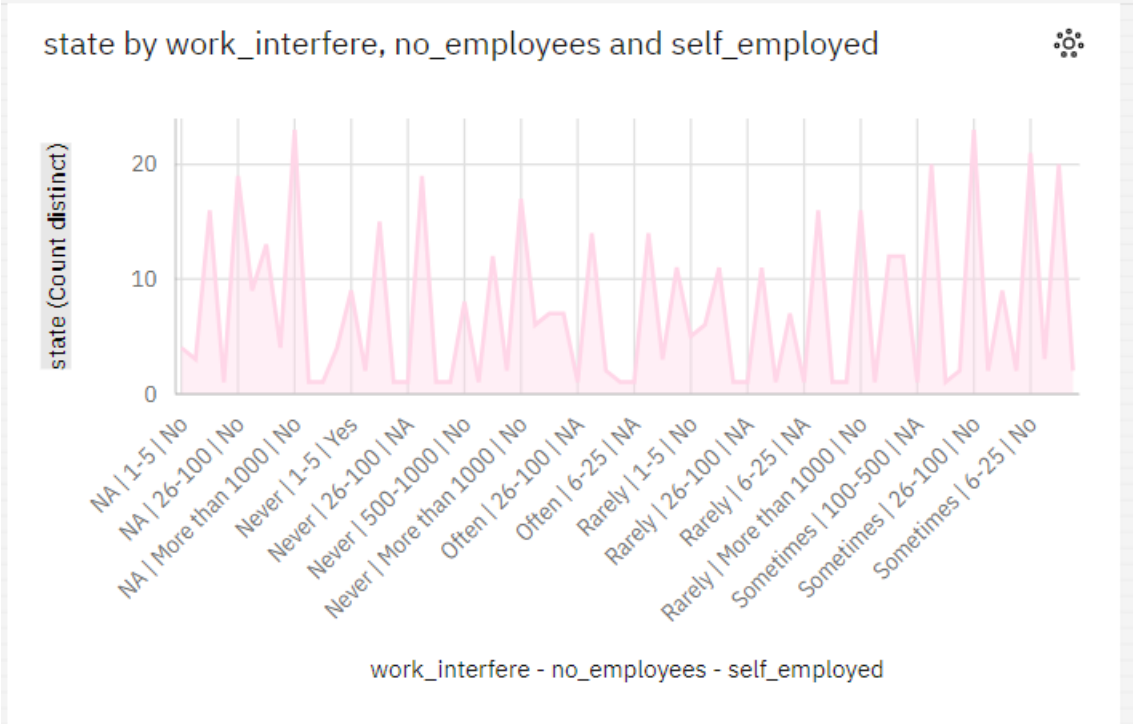
Adapting this plan for healthcare data analysis will help you organize and carry out the necessary steps to analyze healthcare data effectively and create meaningful visualizations using IBM Cognos. Customize the objectives, metrics, and visualizations according to your specific goals and healthcare dataset.


tech_company by no_employees
no_employees
1-5
100-500
26-100
500-1000
6-25
More than 1000
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
tech_company (Count distinct)


self_employed by no_employees
no_employees
6-25
100-500
500-1000
1-5
26-100
More than 1000
NA
No
Yes
self_employed


Gender by self_employed
40
30
20
10
0
Gender (Count distinct)
NA
No
Yes
self_employed


work_interfere and no_employees for self_employed colored by Age
Age (Sum)
596
100,000,004,784
NA
No
Yes
100%
50%
0%
work_interfere (Count distinct)
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
no_employees (Count distinct)


state by work_interfere, no_employees and self_employed
state (Count distinct)
20
10
0
NA | 1-5 | No
NA | 26-100 | No
NA | More than 1000 | No
Never | 1-5 | Yes
Never | 26-100 | NA
Never | 500-1000 | No
Never | More than 1000 | No
Often | 26-100 | NA
Often | 6-25 | NA
Rarely | 1-5 | No
Rarely | 26-100 | NA
Rarely | 6-25 | NA
Rarely | More than 1000 | No
Rarely | More than 1000 | No
Sometimes | 100-500 | NA
Sometimes | 26-100 | No
Sometimes | 6-25 | No
work_interfere - no_employees - self_employed