Dataset 1: Titanic Passenger Data

- **Description**: Details of Titanic passengers, including Survived (0 = No, 1 = Yes), Pclass (1st, 2nd, 3rd), Sex, Age, Fare, Embarked (C = Cherbourg, Q = Queenstown, S = Southampton).
- **Kaggle Link**: https://www.kaggle.com/datasets/c/heloisemartins/titanic-passenger-data ● **File**: `train.csv`

Questions:

- Create a **bar chart** to show the survival rate (%) by passenger class (Pclass).
- Design a **pie chart** to display the proportion of passengers by embarkation port (Embarked).
- Build a **scatter plot** to explore Age vs. Fare, colored by Survived.
- Construct a **stacked bar chart** to compare survival by Sex across Pclass.
- Create a **box plot** to analyze the distribution of Fare by Pclass.
- Develop a **histogram** to show the distribution of passenger ages, split by Survived.

Dataset 2: New York City Airbnb Open Data

- **Description**: Airbnb listings in NYC, including neighborhood, room type (entire home, private room, shared), price, number of reviews, review scores, availability, and property type.
- **Kaggle Link**: https://www.kaggle.com/datasets/dgomonov/new-york-city-airbnb-open-data ● **File**: `AB_NYC_2019.csv`

Questions:

- Create a **bar chart** to show the average price per night by room type.
- Design a **treemap** to display the number of listings by neighborhood group.
- Build a **scatter plot** to explore review scores vs. price, colored by room type.
- Construct a **stacked bar chart** to show the count of listings by property type across neighborhood groups.
- Create a **box plot** to analyze the distribution of prices by neighborhood group.
- Develop a **line chart** to show average availability (Availability 365) by room type.

Dataset 3: Global Video Game Sales

- **Description**: Video game sales data, including game title, platform, year, genre, publisher, and sales (in millions) for North America, Europe, Japan, and globally.
- **Kaggle Link**: https://www.kaggle.com/datasets/gregorut/videogamesales ● **File**: `vgsales.csv`

Questions:

- Create a **bar chart** to show total global sales by genre.
- Design a **pie chart** to display the market share of gaming platforms by global sales.
- Build a **scatter plot** to explore North America sales vs. Europe sales, colored by genre.
- Construct a **stacked bar chart** to compare sales by region (NA, EU, JP) across top 5 publishers.
- Create a **box plot** to analyze the distribution of global sales by platform.
- Develop a **line chart** to show global sales trends over years by genre.

Dataset 4: World Happiness Report

- **Description**: Data on happiness scores for countries, including fields like country, happiness score, GDP per capita, social support, life expectancy, freedom, generosity, and corruption perception.
- **Kaggle Link**: https://www.kaggle.com/datasets/unsdsn/world-happiness
- **File**: `2019.csv` (or latest available year)

Questions:

- Create a **bar chart** to show the top 10 countries by happiness score.
- Design a **choropleth map** to display happiness scores by country.
- Build a **scatter plot** to explore happiness score vs. GDP per capita, colored by region.
- Construct a **stacked bar chart** to compare contributions of GDP, social support, and life expectancy to happiness score for top 5 countries.
- Create a **box plot** to analyze the distribution of happiness scores by region.
- Develop a **heatmap** to show correlations between happiness score, GDP, freedom, and corruption.

Dataset 5: Spotify Tracks Dataset

- **Description**: Data on Spotify tracks, including track name, artist, popularity, duration, genre, danceability, energy, loudness, and other audio features.
- **Kaggle Link**: https://www.kaggle.com/datasets/maharshipandya/spotify-tracks-dataset ● **File**: `dataset.csv`

Questions:

- Create a **bar chart** to show average popularity by genre.
- Design a **pie chart** to display the proportion of tracks by genre.
- Build a **scatter plot** to explore danceability vs. energy, colored by genre.
- Construct a **stacked bar chart** to compare track counts by genre across popularity ranges (e.g., 0-25, 26-50, 51-75, 76-100).
- Create a **box plot** to analyze the distribution of track duration by genre.
- Develop a **histogram** to show the distribution of loudness, split by genre.

Dataset 6: Heart Disease Dataset

- **Description**: Medical data on heart disease, including patient age, sex, chest pain type, resting blood pressure, cholesterol, fasting blood sugar, and target (0 = no heart disease, 1 = heart disease).
- **Kaggle Link**: https://www.kaggle.com/datasets/fedesoriano/heart-disease-dataset ●

File: `heart.csv`

Questions:

- Create a **bar chart** to show the proportion of heart disease cases (target) by sex.
- Design a **pie chart** to display the distribution of chest pain types.
- Build a **scatter plot** to explore age vs. cholesterol, colored by target.
- Construct a **stacked bar chart** to compare heart disease cases by chest pain type across age groups (e.g., <40, 40-60, >60).
- Create a **box plot** to analyze the distribution of resting blood pressure by target.
- Develop a **histogram** to show the distribution of ages, split by target.

Got it — you want to create all these visualizations inside a report in Cognos Analytics!
This changes the approach a little compared to just writing Python code.

What you are trying to do:

- **Upload the datasets (CSV files) into Cognos Analytics**
- **Use Cognos to build:**
  - **Bar charts**
  - **Pie charts**
  - **Scatter plots**
  - **Stacked bar charts**
  - **Box plots**
  - **Histograms**
  - **Treemaps**
  - **Line charts**
  - **Choropleth maps (for happiness report)**
  - **Heatmaps (for happiness correlations)**

General steps for each dataset in Cognos Analytics:

1. **Upload the dataset into Cognos:**
   - **Go to *New* → *Upload Files* → Choose your .csv (e.g., train.csv, AB_NYC_2019.csv, etc.)**
2. **Create a new Report:**
   - **Click *New* → *Report* → *Blank***
3. **Add a Visualization:**
   - **Choose the correct chart type (e.g., Bar, Pie, Scatter, Treemap, etc.)**
   - **Drag and drop fields into:**
     - **Category (X-axis)**
     - **Series (Color)**

        - **Measure (Y-axis or Value)**
4. **Build visualizations one by one for each dataset as per your question list.**
5. **Style the charts (titles, colors, labels) properly.**
6. **Save the report once all charts are added.**

📋 **Here's an example for Dataset 1: Titanic Passenger Data inside Cognos:**

Question	Chart Type	How to Build
Bar chart: Survival rate by Pclass	Bar chart	Use Pclass as Category, average of Survived as Value (multiply by 100 if you want %)
Pie chart: Proportion by Embarked	Pie chart	Use Embarked as Category, count of passengers as Value
Scatter plot: Age vs Fare by Survived	Scatter plot	Age as X-axis, Fare as Y-axis, Survived as color (Series)
Stacked bar: Survival by Sex across Pclass	Stacked bar chart	Pclass as X-axis, Sex as stacked series, count of Survived
Box plot: Fare by Pclass	Box plot	Fare as Value, Pclass as Category
Histogram: Age distribution split by Survived	Histogram	Age as Value, split by Survived

---

📋 **Similarly for Dataset 2: NYC Airbnb:**

Question	Chart Type	How to Build
Average price per night by room type	Bar chart	Room Type as X-axis, average Price as Y-axis
Number of listings by neighborhood group	Treemap	Neighbourhood Group as Category, count of listings
Review scores vs Price by room type	Scatter plot	Review Score as X-axis, Price as Y-axis, Room Type as color
...	...	...

☑ **Special Cases:**

- **Choropleth Map (Happiness Report):**
    - **Use the *Map visualization* in Cognos, set *Country* for location, and *Happiness Score* as the measure.**
- **Heatmap (Happiness correlation):**
    - **Use *Heat Map* chart type.**
    - **Create new calculated fields for correlation if needed.**

📄 **Final Report Structure**

For each dataset:

- **Create a *page or a section* titled (e.g., Titanic Passenger Analysis).**
- **Add all requested charts neatly.**
- **Add short titles like:**
    - **“Survival Rate by Class”**

    - **“Passenger Distribution by Port”**
    - **"Fare Distribution by Class", etc.**
- **Save the report as something like Full Analysis Report.cognosreport**

🛠️ Do you want me to also guide you step-by-step with screenshots or example Cognos expressions if needed?
(For example: how to calculate Survival Rate (%) in Cognos itself?)
👉 Let me know! I can give you a full professional guide 📚 to create this easily! 🚀
Would you like that?