

```
from google.colab import files
uploaded = files.upload()
```

<IPython.core.display.HTML object>

Saving netflix_viewership_dataset.csv to netflix_viewership_dataset (1).csv

```
import pandas as pd
```

```
df = pd.read_csv('netflix_viewership_dataset.csv')
```

```
df=pd.DataFrame(df)
df
```

```
{
  "summary": {
    "name": "df",
    "rows": 100,
    "fields": [
      {
        "column": "Date",
        "properties": {
          "dtype": "object",
          "num_unique_values": 28,
          "samples": [
            "2025-09-19",
            "2025-09-26",
            "2025-09-07"
          ],
          "semantic_type": "",
          "description": ""
        },
        "column": "Region",
        "properties": {
          "dtype": "category",
          "num_unique_values": 8,
          "samples": [
            "Australia",
            "USA",
            "India"
          ],
          "semantic_type": "",
          "description": ""
        },
        "column": "Movie Title",
        "properties": {
          "dtype": "category",
          "num_unique_values": 15,
          "samples": [
            "Triple Frontier",
            "Bird Box",
            "Don't Look Up"
          ],
          "semantic_type": "",
          "description": ""
        },
        "column": "Genre",
        "properties": {
          "dtype": "category",
          "num_unique_values": 7,
          "samples": [
            "Drama",
            "Thriller",
            "Documentary"
          ],
          "semantic_type": "",
          "description": ""
        },
        "column": "Views (millions)",
        "properties": {
          "dtype": "number",
          "std": 4.1969949266280295,
          "min": 0.6,
          "max": 15.0,
          "num_unique_values": 98,
          "samples": [
            8.0,
            6.64,
            10.22
          ],
          "semantic_type": "",
          "description": ""
        },
        "column": "Watch Time (hours)",
        "properties": {
          "dtype": "number",
          "std": 12.333858748431595,
          "min": 1.18,
          "max": 49.83,
          "num_unique_values": 97,
          "samples": [
            14.76,
            5.94,
            23.83
          ],
          "semantic_type": "",
          "description": ""
        },
        "column": "Revenue (USD millions)",
        "properties": {
          "dtype": "number",
          "std": 6.458923714168993,
          "min": 0.92,
          "max": 27.8,
          "num_unique_values": 97,
          "samples": [
            14.76,
            5.94,
            23.83
          ],
          "semantic_type": "",
          "description": ""
        }
      ]
    }
  }
}
```

```

\"num_unique_values\": 93,\n          \"samples\": [\n          2.82,\n          5.59,\n          8.54\n          ],\n          \"semantic_type\": \"\",\n          \"description\": \"\"\n        }\n      },\n      {\n        \"column\":\n        \"Subscriber Type\",\n        \"properties\": {\n          \"dtype\":\n          \"category\",\n          \"num_unique_values\": 3,\n          \"samples\":\n          [\n            \"Standard\",\n            \"Premium\",\n            \"Basic\"\n          ],\n          \"semantic_type\": \"\",\n          \"description\": \"\"\n        }\n      }\n    ]\n  },\n  \"type\": \"dataframe\", \"variable_name\": \"df\"}

```

```
df.head()
```

```

{"summary": "{\n  \"name\": \"df\",\n  \"rows\": 100,\n  \"fields\": [\n    {\n      \"column\": \"Date\",\n      \"properties\": {\n        \"dtype\": \"object\",\n        \"num_unique_values\": 28,\n        \"samples\": [\n          \"2025-09-19\",\n          \"2025-09-26\",\n          \"2025-09-07\"\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      },\n      {\n        \"column\":\n        \"Region\",\n        \"properties\": {\n          \"dtype\":\n          \"category\",\n          \"num_unique_values\": 8,\n          \"samples\":\n          [\n            \"Australia\",\n            \"USA\",\n            \"India\"\n          ],\n          \"semantic_type\": \"\",\n          \"description\": \"\"\n        }\n      },\n      {\n        \"column\": \"Movie Title\",\n        \"properties\": {\n          \"dtype\": \"category\",\n          \"num_unique_values\": 15,\n          \"samples\": [\n            \"Triple Frontier\",\n            \"Bird Box\",\n            \"Don't Look Up\"\n          ],\n          \"semantic_type\": \"\",\n          \"description\": \"\"\n        }\n      },\n      {\n        \"column\": \"Genre\",\n        \"properties\": {\n          \"dtype\": \"category\",\n          \"num_unique_values\": 7,\n          \"samples\": [\n            \"Drama\",\n            \"Thriller\",\n            \"Documentary\"\n          ],\n          \"semantic_type\": \"\",\n          \"description\": \"\"\n        }\n      },\n      {\n        \"column\": \"Views (millions)\",\n        \"properties\": {\n          \"dtype\": \"number\",\n          \"std\": 4.1969949266280295,\n          \"min\": 0.6,\n          \"max\": 15.0,\n          \"num_unique_values\": 98,\n          \"samples\": [\n            8.0,\n            6.64,\n            10.22\n          ],\n          \"semantic_type\": \"\",\n          \"description\": \"\"\n        }\n      },\n      {\n        \"column\":\n        \"Watch Time (hours)\",\n        \"properties\": {\n          \"dtype\":\n          \"number\",\n          \"std\": 12.333858748431595,\n          \"min\": 1.18,\n          \"max\": 49.83,\n          \"num_unique_values\": 97,\n          \"samples\": [\n            14.76,\n            5.94,\n            23.83\n          ],\n          \"semantic_type\": \"\",\n          \"description\": \"\"\n        }\n      },\n      {\n        \"column\": \"Revenue (USD millions)\",\n        \"properties\": {\n          \"dtype\": \"number\",\n          \"std\": 6.458923714168993,\n          \"min\": 0.92,\n          \"max\": 27.8,\n          \"num_unique_values\": 93,\n          \"samples\": [\n            2.82,\n            5.59,\n            8.54\n          ],\n          \"semantic_type\": \"\",\n          \"description\": \"\"\n        }\n      },\n      {\n        \"column\":\n        \"Subscriber Type\",\n        \"properties\": {\n          \"dtype\":

```

```
"category","\n          \"num_unique_values\": 3,\n          \"samples\": [\n            \"Standard\",\n            \"Premium\",\n            \"Basic\"],\n          \"semantic_type\": \"\", \n          \"description\": \"\"\n        }\n      ]\n    },\n    \"type\": \"dataframe\", \"variable_name\": \"df\"}
```

```
df.tail()
```

```
{"repr_error": "0", "type": "dataframe"}
```

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
```

RangeIndex: 100 entries, 0 to 99

Data columns (total 8 columns):

#	Column	Non-Null Count		Dtype
0	Date	100	non-null	object
1	Region	100	non-null	object
2	Movie Title	100	non-null	object
3	Genre	100	non-null	object
4	Views (millions)	100	non-null	float64
5	Watch Time (hours)	100	non-null	float64
6	Revenue (USD millions)	100	non-null	float64
7	Subscriber Type	100	non-null	object

```
dtypes: float64(3), object(5)
```

```
memory usage: 6.4+ KB
```

```
df.describe()
```

```
{
  "summary": "{\n  \"name\": \"df\",\n  \"rows\": 8,\n  \"fields\": [\n    {\n      \"column\": \"Views (millions)\",\n      \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 33.14862816036609,\n        \"min\": 0.6,\n        \"max\": 100.0,\n        \"num_unique_values\": 8,\n        \"samples\": [\n          7.220000000000001,\n          100.0\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      }\n    },\n    {\n      \"column\": \"Watch Time (hours)\",\n      \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 31.868588385387344,\n        \"min\": 1.18,\n        \"max\": 100.0,\n        \"num_unique_values\": 8,\n        \"samples\": [\n          17.22,\n          100.0\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      }\n    },\n    {\n      \"column\": \"Revenue (USD millions)\",\n      \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 32.69020944575808,\n        \"min\": 0.92,\n        \"max\": 100.0,\n        \"num_unique_values\": 8,\n        \"samples\": [\n          9.982099999999999,\n          8.395,\n          100.0\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      }\n    }\n  ]\n}",
  "type": "dataframe"
}
```

```

df.shape
(100, 8)

df.columns
Index(['Date', 'Region', 'Movie Title', 'Genre', 'Views (millions)',
      'Watch Time (hours)', 'Revenue (USD millions)', 'Subscriber
      Type'],
      dtype='object')

df.values
array([[ '2025-09-06', 'India', "Don't Look Up", 'Drama', 1.88, 4.58,
        3.11, 'Standard'],
       [ '2025-09-29', 'Australia', 'Extraction 2', 'Thriller', 12.51,
        21.21, 13.72, 'Standard'],
       [ '2025-09-28', 'India', 'Project Power', 'Drama', 10.47, 29.51,
        16.69, 'Standard'],
       [ '2025-09-24', 'UK', 'The Adam Project', 'Horror', 7.35, 19.89,
        6.83, 'Standard'],
       [ '2025-09-09', 'Australia', 'The Old Guard', 'Thriller', 2.22,
        6.69, 3.49, 'Premium'],
       [ '2025-09-30', 'Germany', 'Red Notice', 'Romance', 9.16, 17.72,
        15.82, 'Basic'],
       [ '2025-09-23', 'Australia', 'Army of the Dead', 'Drama', 5.15,
        13.62, 8.19, 'Basic'],
       [ '2025-09-20', 'Japan', 'The Gray Man', 'Horror', 4.71, 12.8,
        7.72, 'Standard'],
       [ '2025-09-23', 'USA', 'The Adam Project', 'Thriller', 4.32,
        9.6, 4.56, 'Premium'],
       [ '2025-09-07', 'Germany', 'The Gray Man', 'Action', 6.83,
        18.44, 11.56, 'Standard'],
       [ '2025-09-19', 'Brazil', 'Heart of Stone', 'Documentary', 8.23,
        26.04, 9.78, 'Standard'],
       [ '2025-09-20', 'India', 'Triple Frontier', 'Action', 5.17,
        9.08, 5.69, 'Standard'],
       [ '2025-09-18', 'UK', 'Heart of Stone', 'Drama', 5.14, 13.16,
        7.33, 'Premium'],
       [ '2025-09-18', 'Australia', 'Glass Onion', 'Action', 10.18,
        32.9, 10.85, 'Premium'],
       [ '2025-09-21', 'Germany', 'Triple Frontier', 'Horror', 0.89,
        1.98, 1.05, 'Premium'],
       [ '2025-09-28', 'Canada', 'Bird Box', 'Thriller', 8.69, 13.74,

```

8.3,
 'Standard'],
 ['2025-09-15', 'Japan', 'Glass Onion', 'Action', 2.31, 7.87,
 1.94,
 'Standard'],
 ['2025-09-12', 'Japan', "Don't Look Up", 'Documentary', 3.33,
 11.1, 4.82, 'Premium'],
 ['2025-09-02', 'Japan', 'Project Power', 'Thriller', 7.64,
 19.29,
 6.15, 'Premium'],
 ['2025-09-24', 'UK', 'Extraction 2', 'Comedy', 13.25, 40.69,
 23.47, 'Standard'],
 ['2025-09-27', 'Japan', 'Heart of Stone', 'Comedy', 8.48,
 20.63,
 15.82, 'Basic'],
 ['2025-09-01', 'Germany', 'Bird Box', 'Action', 5.06, 17.7,
 4.88,
 'Standard'],
 ['2025-09-29', 'UK', 'Heart of Stone', 'Horror', 5.93, 13.25,
 4.9,
 'Premium'],
 ['2025-09-18', 'Australia', 'Triple Frontier', 'Horror', 3.63,
 8.9, 5.59, 'Standard'],
 ['2025-09-06', 'India', 'The Adam Project', 'Romance', 3.99,
 7.51,
 4.41, 'Premium'],
 ['2025-09-01', 'Australia', 'Heart of Stone', 'Drama', 2.13,
 3.79,
 1.98, 'Standard'],
 ['2025-09-30', 'India', 'Bird Box', 'Thriller', 1.24, 3.08,
 1.9,
 'Basic'],
 ['2025-09-08', 'Japan', 'Army of the Dead', 'Horror', 4.64,
 15.21,
 8.05, 'Premium'],
 ['2025-09-06', 'India', 'Red Notice', 'Romance', 11.34, 30.36,
 9.78, 'Premium'],
 ['2025-09-03', 'Australia', 'Red Notice', 'Horror', 5.03,
 12.65,
 8.7, 'Standard'],
 ['2025-09-07', 'USA', 'The Old Guard', 'Drama', 12.83, 41.03,
 22.13, 'Basic'],
 ['2025-09-23', 'Japan', 'The Mother', 'Documentary', 5.31,
 9.73,
 5.75, 'Standard'],
 ['2025-09-08', 'UK', 'Army of the Dead', 'Action', 8.59, 25.76,
 11.47, 'Premium'],
 ['2025-09-11', 'Japan', 'The Mother', 'Thriller', 15.0, 49.83,
 23.46, 'Basic'],

['2025-09-07', 'Germany', 'Enola Holmes 2', 'Comedy', 14.74, 28.15, 27.8, 'Basic'],
 ['2025-09-10', 'Japan', 'Army of the Dead', 'Comedy', 1.61, 3.67, 2.41, 'Premium'],
 ['2025-09-28', 'Australia', 'Red Notice', 'Horror', 11.61, 18.2, 16.54, 'Premium'],
 ['2025-09-19', 'Canada', 'Extraction 2', 'Action', 0.67, 1.18, 1.0, 'Basic'],
 ['2025-09-16', 'Australia', 'Project Power', 'Comedy', 6.93, 13.92, 7.43, 'Basic'],
 ['2025-09-03', 'USA', 'Red Notice', 'Romance', 11.09, 37.33, 15.99, 'Premium'],
 ['2025-09-13', 'Japan', 'The Old Guard', 'Thriller', 1.88, 5.94, 1.62, 'Basic'],
 ['2025-09-09', 'Japan', 'Project Power', 'Romance', 6.64, 23.04, 10.66, 'Standard'],
 ['2025-09-05', 'UK', 'Triple Frontier', 'Action', 9.53, 32.12, 15.82, 'Premium'],
 ['2025-09-13', 'India', 'Army of the Dead', 'Action', 3.2, 11.18, 2.82, 'Premium'],
 ['2025-09-29', 'Australia', 'Bird Box', 'Action', 4.85, 9.23, 8.49, 'Premium'],
 ['2025-09-14', 'Australia', 'Extraction 2', 'Romance', 14.61, 27.78, 26.52, 'Standard'],
 ['2025-09-15', 'Brazil', 'Heart of Stone', 'Horror', 2.97, 6.2, 3.69, 'Basic'],
 ['2025-09-26', 'Australia', 'The Gray Man', 'Action', 13.37, 41.03, 16.82, 'Basic'],
 ['2025-09-26', 'UK', 'Bird Box', 'Thriller', 1.69, 4.95, 2.22, 'Basic'],
 ['2025-09-30', 'Brazil', 'Extraction 2', 'Action', 14.53, 46.69, 15.99, 'Standard'],
 ['2025-09-29', 'UK', 'Enola Holmes 2', 'Thriller', 11.89, 40.69, 11.97, 'Basic'],
 ['2025-09-30', 'Germany', 'The Old Guard', 'Comedy', 12.46, 42.63, 10.24, 'Standard'],
 ['2025-09-30', 'USA', "Don't Look Up", 'Romance', 5.13, 16.54, 4.4, 'Basic'],
 ['2025-09-14', 'USA', 'Project Power', 'Romance', 14.75, 25.3, 19.17, 'Basic'],
 ['2025-09-19', 'Australia', 'Spaceman', 'Horror', 8.46, 25.77,

12.5, 'Basic'],
 ['2025-09-04', 'Canada', 'The Gray Man', 'Thriller', 9.77,
 33.48,
 14.34, 'Basic'],
 ['2025-09-11', 'USA', 'Spaceman', 'Thriller', 2.83, 9.32, 5.31,
 'Standard'],
 ['2025-09-13', 'Canada', 'Enola Holmes 2', 'Drama', 6.2, 17.21,
 6.29, 'Standard'],
 ['2025-09-30', 'India', 'The Mother', 'Romance', 8.68, 21.58,
 14.26, 'Standard'],
 ['2025-09-13', 'Canada', "Don't Look Up", 'Drama', 9.9, 24.59,
 8.54, 'Premium'],
 ['2025-09-20', 'USA', 'Spaceman', 'Comedy', 1.32, 2.7, 2.29,
 'Standard'],
 ['2025-09-19', 'Japan', 'Red Notice', 'Thriller', 8.33, 23.99,
 8.11, 'Basic'],
 ['2025-09-04', 'UK', 'Enola Holmes 2', 'Action', 4.28, 14.62,
 4.11, 'Premium'],
 ['2025-09-02', 'Australia', 'The Old Guard', 'Romance', 8.0,
 12.36, 13.91, 'Basic'],
 ['2025-09-27', 'Japan', 'Enola Holmes 2', 'Drama', 8.93, 14.76,
 9.14, 'Basic'],
 ['2025-09-17', 'USA', 'The Old Guard', 'Comedy', 14.82, 43.78,
 18.39, 'Premium'],
 ['2025-09-23', 'Canada', 'The Adam Project', 'Comedy', 6.36,
 16.98, 6.74, 'Standard'],
 ['2025-09-12', 'Japan', 'The Gray Man', 'Horror', 10.49, 22.35,
 13.39, 'Basic'],
 ['2025-09-09', 'Germany', 'Project Power', 'Comedy', 5.91,
 16.77,
 5.37, 'Basic'],
 ['2025-09-13', 'Brazil', 'Triple Frontier', 'Action', 4.53,
 15.03,
 4.19, 'Standard'],
 ['2025-09-08', 'India', 'Red Notice', 'Drama', 11.69, 29.11,
 23.05, 'Premium'],
 ['2025-09-24', 'Japan', 'The Adam Project', 'Horror', 5.07,
 9.97,
 7.88, 'Basic'],
 ['2025-09-18', 'UK', 'Enola Holmes 2', 'Romance', 1.22, 2.83,
 2.18, 'Premium'],
 ['2025-09-23', 'India', 'Bird Box', 'Documentary', 8.65, 22.0,
 15.03, 'Basic'],
 ['2025-09-13', 'Germany', 'Army of the Dead', 'Thriller', 3.81,
 11.85, 5.93, 'Standard'],
 ['2025-09-08', 'Brazil', "Don't Look Up", 'Drama', 1.24, 4.02,
 2.34, 'Premium'],
 ['2025-09-29', 'Brazil', 'The Gray Man', 'Thriller', 0.79,
 1.77,

0.92, 'Basic'],
 ['2025-09-06', 'USA', 'The Old Guard', 'Romance', 9.47, 27.99,
 17.98, 'Premium'],
 ['2025-09-23', 'India', 'Project Power', 'Comedy', 7.41, 17.23,
 6.96, 'Basic'],
 ['2025-09-15', 'Canada', 'Enola Holmes 2', 'Drama', 14.83,
 33.1,
 19.12, 'Basic'],
 ['2025-09-06', 'Brazil', 'Red Notice', 'Thriller', 10.86,
 36.84,
 12.98, 'Premium'],
 ['2025-09-28', 'Germany', 'Triple Frontier', 'Comedy', 6.16,
 19.72, 5.5, 'Basic'],
 ['2025-09-11', 'Canada', 'Heart of Stone', 'Action', 4.04,
 9.89,
 3.91, 'Basic'],
 ['2025-09-19', 'Australia', 'Red Notice', 'Documentary', 5.73,
 16.58, 7.52, 'Premium'],
 ['2025-09-17', 'Japan', 'Bird Box', 'Comedy', 4.87, 15.52,
 7.52,
 'Standard'],
 ['2025-09-24', 'UK', "Don't Look Up", 'Comedy', 0.6, 2.05,
 0.92,
 'Premium'],
 ['2025-09-27', 'Germany', 'Heart of Stone', 'Documentary',
 11.04,
 32.94, 15.55, 'Standard'],
 ['2025-09-27', 'Canada', 'Bird Box', 'Thriller', 12.66, 34.04,
 17.18, 'Standard'],
 ['2025-09-12', 'UK', 'Triple Frontier', 'Action', 14.5, 38.96,
 19.41, 'Basic'],
 ['2025-09-24', 'Canada', 'The Adam Project', 'Horror', 13.64,
 45.79, 20.29, 'Basic'],
 ['2025-09-12', 'Canada', 'Enola Holmes 2', 'Drama', 9.94,
 16.42,
 13.4, 'Premium'],
 ['2025-09-06', 'Brazil', 'The Old Guard', 'Drama', 4.01, 8.14,
 7.97, 'Basic'],
 ['2025-09-01', 'India', 'The Gray Man', 'Romance', 7.19, 20.15,
 9.49, 'Standard'],
 ['2025-09-19', 'UK', 'Triple Frontier', 'Horror', 7.25, 19.28,
 8.53, 'Premium'],
 ['2025-09-03', 'UK', 'Army of the Dead', 'Action', 11.15,
 17.15,
 10.22, 'Standard'],
 ['2025-09-23', 'Australia', 'The Old Guard', 'Comedy', 6.94,
 23.83, 5.93, 'Standard'],
 ['2025-09-04', 'Canada', 'The Old Guard', 'Action', 10.22,
 25.76,

```

9.02, 'Premium'],
['2025-09-20', 'Canada', 'Glass Onion', 'Action', 14.9, 34.36,
18.6, 'Basic'],
['2025-09-14', 'UK', 'The Adam Project', 'Drama', 14.94, 49.58,
21.41, 'Standard'],
['2025-09-30', 'Germany', 'The Adam Project', 'Romance', 8.24,
14.26, 13.1, 'Basic']], dtype=object)

```

```
df.dtypes
```

```

Date                object
Region              object
Movie Title         object
Genre               object
Views (millions)    float64
Watch Time (hours)  float64
Revenue (USD millions) float64
Subscriber Type     object
dtype: object

```

```
df.loc[0]
```

```

Date                2025-09-06
Region              India
Movie Title         Don't Look Up
Genre               Drama
Views (millions)    1.88
Watch Time (hours)  4.58
Revenue (USD millions) 3.11
Subscriber Type     Standard
Name: 0, dtype: object

```

```

df_dropped=df.drop(columns=['Revenue (USD millions)'])
df_dropped

```

```

{"summary":{"name": "df_dropped", "rows": 100,
"fields": [{"column": "Date",
"properties": {"dtype": "object",
"num_unique_values": 28,
"samples": ["2025-09-19",
"2025-09-26",
"2025-09-07",
],
"semantic_type": "",
"description": ""}
}, {"column": "Region",
"properties": {"dtype": "category",
"num_unique_values": 8,
"samples": ["Australia",
"USA",
"India",
],
"semantic_type": "",
"description": ""}
}, {"column": "Movie Title",
"properties": {"dtype": "category",
"num_unique_values": 15,
"samples": ["Triple Frontier",
"Box",
"Don't Look Up",
],
"semantic_type": "",
"description": ""}
}
}

```

```

n    },\n    {\n        \"column\": \"Genre\", \n        \"properties\": {\n            \"dtype\": \"category\", \n            \"num_unique_values\": 7, \n            \"samples\": [\n                \"Drama\", \n                \"Thriller\", \n                \"Documentary\" \n            ], \n            \"semantic_type\": \"\", \n            \"description\": \"\" \n        } \n    }, \n    {\n        \"column\": \"Views (millions)\", \n        \"properties\": {\n            \"dtype\": \"number\", \n            \"std\": 4.1969949266280295, \n            \"min\": 0.6, \n            \"max\": 15.0, \n            \"num_unique_values\": 98, \n            \"samples\": [\n                8.0, \n                6.64, \n                10.22 \n            ], \n            \"semantic_type\": \"\", \n            \"description\": \"\" \n        } \n    }, \n    {\n        \"column\": \"Watch Time (hours)\", \n        \"properties\": {\n            \"dtype\": \"number\", \n            \"std\": 12.333858748431595, \n            \"min\": 1.18, \n            \"max\": 49.83, \n            \"num_unique_values\": 97, \n            \"samples\": [\n                14.76, \n                5.94, \n                23.83 \n            ], \n            \"semantic_type\": \"\", \n            \"description\": \"\" \n        } \n    }, \n    {\n        \"column\": \"Subscriber Type\", \n        \"properties\": {\n            \"dtype\": \"category\", \n            \"num_unique_values\": 3, \n            \"samples\": [\n                \"Standard\", \n                \"Premium\", \n                \"Basic\" \n            ], \n            \"semantic_type\": \"\", \n            \"description\": \"\" \n        } \n    } \n]\n} \", \"type\": \"dataframe\", \"variable_name\": \"df_dropped\"}

```

```

df_renamed = df.rename(columns={'Movie Title': 'Movie'})
df_renamed

```

```

{"summary": "{\n    \"name\": \"df_renamed\", \n    \"rows\": 100, \n    \"fields\": [\n        {\n            \"column\": \"Date\", \n            \"properties\": {\n                \"dtype\": \"object\", \n                \"num_unique_values\": 28, \n                \"samples\": [\n                    \"2025-09-19\", \n                    \"2025-09-26\", \n                    \"2025-09-07\" \n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\" \n            } \n        }, \n        {\n            \"column\": \"Region\", \n            \"properties\": {\n                \"dtype\": \"category\", \n                \"num_unique_values\": 8, \n                \"samples\": [\n                    \"Australia\", \n                    \"USA\", \n                    \"India\" \n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\" \n            } \n        }, \n        {\n            \"column\": \"Movie\", \n            \"properties\": {\n                \"dtype\": \"category\", \n                \"num_unique_values\": 15, \n                \"samples\": [\n                    \"Triple Frontier\", \n                    \"Bird Box\", \n                    \"Don't Look Up\" \n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\" \n            } \n        }, \n        {\n            \"column\": \"Genre\", \n            \"properties\": {\n                \"dtype\": \"category\", \n                \"num_unique_values\": 7, \n                \"samples\": [\n                    \"Drama\", \n                    \"Thriller\", \n                    \"Documentary\" \n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\" \n            } \n        }, \n        {\n            \"column\": \"Views (millions)\", \n            \"properties\": {\n                \"dtype\": \"number\", \n                \"std\": 4.1969949266280295, \n                \"min\": 0.6, \n                \"max\": 15.0, \n                \"num_unique_values\": 98, \n            } \n        } \n    ] \n} \", \"type\": \"dataframe\", \"variable_name\": \"df_renamed\"}

```



```

}\n    },\n    {\n        \"column\": \"Watch Time (hours)\",\n        \"properties\": {\n            \"dtype\": \"number\", \n            \"std\": 12.333858748431599,\n            \"min\": 1.18,\n            \"max\": 49.83,\n            \"num_unique_values\": 97,\n            \"samples\": [\n                21.58,\n                16.77,\n                33.1\n            ],\n            \"semantic_type\": \"\",\n            \"description\": \"\"\n        },\n        {\n            \"column\": \"Revenue (USD millions)\",\n            \"properties\": {\n                \"dtype\": \"number\", \n                \"std\": 6.45892371416899,\n                \"min\": 0.92,\n                \"max\": 27.8,\n                \"num_unique_values\": 93,\n                \"samples\": [\n                    5.5,\n                    3.91,\n                    12.5\n                ],\n                \"semantic_type\": \"\",\n                \"description\": \"\"\n            },\n            {\n                \"column\": \"Subscriber Type\",\n                \"properties\": {\n                    \"dtype\": \"category\", \n                    \"num_unique_values\": 3,\n                    \"samples\": [\n                        \"Premium\", \n                        \"Basic\", \n                        \"Standard\"\n                    ],\n                    \"semantic_type\": \"\",\n                    \"description\": \"\"\n                }\n            }\n        ],\n        \"type\": \"dataframe\", \"variable_name\": \"df_sorted\"}

```

```

df_filled = df.fillna(0)
df_filled

```

```

{"summary": "{\n    \"name\": \"df_filled\", \n    \"rows\": 100,\n    \"fields\": [\n        {\n            \"column\": \"Date\", \n            \"properties\": {\n                \"dtype\": \"object\", \n                \"num_unique_values\": 28,\n                \"samples\": [\n                    \"2025-09-19\", \n                    \"2025-09-26\", \n                    \"2025-09-07\"\n                ],\n                \"semantic_type\": \"\", \n                \"description\": \"\"\n            }\n        },\n        {\n            \"column\": \"Region\", \n            \"properties\": {\n                \"dtype\": \"category\", \n                \"num_unique_values\": 8,\n                \"samples\": [\n                    \"Australia\", \n                    \"USA\", \n                    \"India\"\n                ],\n                \"semantic_type\": \"\", \n                \"description\": \"\"\n            }\n        },\n        {\n            \"column\": \"Movie Title\", \n            \"properties\": {\n                \"dtype\": \"category\", \n                \"num_unique_values\": 15,\n                \"samples\": [\n                    \"Triple Frontier\", \n                    \"Bird Box\", \n                    \"Don't Look Up\"\n                ],\n                \"semantic_type\": \"\", \n                \"description\": \"\"\n            }\n        },\n        {\n            \"column\": \"Genre\", \n            \"properties\": {\n                \"dtype\": \"category\", \n                \"num_unique_values\": 7,\n                \"samples\": [\n                    \"Drama\", \n                    \"Thriller\", \n                    \"Documentary\"\n                ],\n                \"semantic_type\": \"\", \n                \"description\": \"\"\n            }\n        },\n        {\n            \"column\": \"Views (millions)\", \n            \"properties\": {\n                \"dtype\": \"number\", \n                \"std\": 4.1969949266280295,\n                \"min\": 0.6,\n                \"max\": 15.0,\n                \"num_unique_values\": 98,\n                \"samples\": [\n                    8.0,\n                    6.64,\n                    10.22\n                ],\n                \"semantic_type\": \"\", \n                \"description\": \"\"\n            }\n        },\n        {\n            \"column\": \"Watch Time (hours)\", \n            \"properties\": {\n                \"dtype\": \"number\", \n                \"std\":

```

```

12.333858748431595,\n          \"min\": 1.18,\n          \"max\": 49.83,\n          \"num_unique_values\": 97,\n          \"samples\": [\n          14.76,\n          5.94,\n          23.83\n          ],\n          \"semantic_type\": \"\",\n          \"description\": \"\"\n        }\n      },\n      {\n        \"column\":\n        \"Revenue (USD millions)\",\n        \"properties\": {\n          \"dtype\": \"number\",\n          \"std\": 6.458923714168993,\n          \"min\": 0.92,\n          \"max\": 27.8,\n          \"num_unique_values\":\n          93,\n          \"samples\": [\n          2.82,\n          5.59,\n          8.54\n          ],\n          \"semantic_type\": \"\",\n          \"description\": \"\"\n        },\n        {\n          \"column\":\n          \"Subscriber Type\",\n          \"properties\": {\n            \"dtype\":\n            \"category\",\n            \"num_unique_values\": 3,\n            \"samples\":\n            [\n              \"Standard\",\n              \"Premium\",\n              \"Basic\"\n            ],\n            \"semantic_type\": \"\",\n            \"description\": \"\"\n          }\n        }\n      }\n    ]\n  },\n  \"type\": \"dataframe\", \"variable_name\": \"df_filled\"}

```

```

df_unique = df.drop_duplicates()
df_unique

```

```

{"summary": "{\n  \"name\": \"df_unique\",\n  \"rows\": 100,\n  \"fields\": [\n    {\n      \"column\": \"Date\",\n      \"properties\": {\n        \"dtype\": \"object\",\n        \"num_unique_values\": 28,\n        \"samples\": [\n          \"2025-09-19\",\n          \"2025-09-26\",\n          \"2025-09-07\"\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      },\n      {\n        \"column\": \"Region\",\n        \"properties\": {\n          \"dtype\": \"category\",\n          \"num_unique_values\":\n          8,\n          \"samples\": [\n            \"Australia\",\n            \"USA\",\n            \"India\"\n          ],\n          \"semantic_type\":\n          \"\",\n          \"description\": \"\"\n        },\n        {\n          \"column\": \"Movie Title\",\n          \"properties\": {\n            \"dtype\": \"category\",\n            \"num_unique_values\": 15,\n            \"samples\": [\n              \"Triple Frontier\",\n              \"Bird Box\",\n              \"Don't Look Up\"\n            ],\n            \"semantic_type\": \"\",\n            \"description\": \"\"\n          },\n          {\n            \"column\": \"Genre\",\n            \"properties\": {\n              \"dtype\": \"category\",\n              \"num_unique_values\": 7,\n              \"samples\": [\n                \"Drama\",\n                \"Thriller\",\n                \"Documentary\"\n              ],\n              \"semantic_type\": \"\",\n              \"description\": \"\"\n            },\n            {\n              \"column\":\n              \"Views (millions)\",\n              \"properties\": {\n                \"dtype\":\n                \"number\",\n                \"std\": 4.1969949266280295,\n                \"min\":\n                0.6,\n                \"max\": 15.0,\n                \"num_unique_values\": 98,\n                \"samples\": [\n                  8.0,\n                  6.64,\n                  10.22\n                ],\n                \"semantic_type\": \"\",\n                \"description\": \"\"\n              },\n              {\n                \"column\":\n                \"Watch Time (hours)\",\n                \"properties\": {\n                  \"dtype\": \"number\",\n                  \"std\":\n                  12.333858748431595,\n                  \"min\": 1.18,\n                  \"max\": 49.83,\n                  \"num_unique_values\": 97,\n                  \"samples\": [\n                    14.76,\n                    5.94,\n                    23.83\n                    ],\n                    \"semantic_type\": \"\",\n                    \"description\": \"\"\n                  }\n                }\n              }\n            }\n          }\n        }\n      }\n    ]\n  },\n  \"type\": \"dataframe\", \"variable_name\": \"df_unique\"}

```

```
5.94,\n                23.83\n            ],\n            \"semantic_type\": \"\",\n            \"description\": \"\"\n        },\n        {\n            \"column\":\n            \"Revenue (USD millions)\",\n            \"properties\": {\n                \"dtype\": \"number\",\n                \"std\": 6.458923714168993,\n                \"min\": 0.92,\n                \"max\": 27.8,\n                \"num_unique_values\":\n                93,\n                \"samples\": [\n                    2.82,\n                    5.59,\n                    8.54\n                ],\n                \"semantic_type\": \"\",\n                \"description\": \"\"\n            },\n            {\n                \"column\":\n                \"Subscriber Type\",\n                \"properties\": {\n                    \"dtype\":\n                    \"category\",\n                    \"num_unique_values\": 3,\n                    \"samples\":\n                    [\n                        \"Standard\",\n                        \"Premium\",\n                        \"Basic\"\n                    ],\n                    \"semantic_type\": \"\",\n                    \"description\": \"\"\n                }\n            }\n        ]\n    },\n    \"type\": \"dataframe\", \"variable_name\": \"df_unique\"}
```

```
df_replaced = df.replace({'Project Power': 'Power puff'})
df_replaced
```

```
{
  "summary": {
    "name": "df_replaced",
    "rows": 100,
    "fields": [
      {
        "column": "Date",
        "dtype": "object",
        "num_unique_values": 28,
        "samples": [
          "2025-09-19",
          "2025-09-26",
          "2025-09-07"
        ],
        "semantic_type": "",
        "description": ""
      },
      {
        "column": "Region",
        "dtype": "category",
        "num_unique_values": 8,
        "samples": [
          "Australia",
          "USA",
          "India"
        ],
        "semantic_type": "",
        "description": ""
      },
      {
        "column": "Movie Title",
        "dtype": "category",
        "num_unique_values": 15,
        "samples": [
          "Triple Frontier",
          "Bird Box",
          "Don't Look Up"
        ],
        "semantic_type": "",
        "description": ""
      },
      {
        "column": "Genre",
        "dtype": "category",
        "num_unique_values": 7,
        "samples": [
          "Drama",
          "Thriller",
          "Documentary"
        ],
        "semantic_type": "",
        "description": ""
      },
      {
        "column": "Views (millions)",
        "dtype": "number",
        "std": 4.1969949266280295,
        "min": 0.6,
        "max": 15.0,
        "num_unique_values": 98,
        "samples": [
          8.0,
          6.64,
          10.22
        ],
        "semantic_type": "",
        "description": ""
      },
      {
        "column": "Watch Time (hours)",
        "dtype": "number",
        "std": 12.333858748431595,
        "min": 1.18,
        "max": 49.83,
        "num_unique_values": 97,
        "samples": [
          14.76,
          5.94,
          23.83
        ],
        "semantic_type": "",
        "description": ""
      }
    ]
  }
}
```

```

\"Revenue (USD millions)\",\n        \"properties\": {\n
\"dtype\": \"number\", \n        \"std\": 6.458923714168993,\n
\"min\": 0.92,\n        \"max\": 27.8,\n        \"num_unique_values\":\n
93,\n        \"samples\": [\n        2.82,\n        5.59,\n
8.54\n        ],\n        \"semantic_type\": \"\", \n
\"description\": \"\"\n        }\n    },\n    {\n        \"column\":\n
\"Subscriber Type\", \n        \"properties\": {\n        \"dtype\":\n
\"category\", \n        \"num_unique_values\": 3,\n        \"samples\":\n
[\n        \"Standard\", \n        \"Premium\", \n
\"Basic\"\n        ],\n        \"semantic_type\": \"\", \n
\"description\": \"\"\n        }\n    }\n    ]\n
n} \", \"type\": \"dataframe\", \"variable_name\": \"df_replaced\"}

```

```

grouped_df = df.groupby('Movie Title').sum()
grouped_df

```

```

{\"summary\": {\"name\": \"grouped_df\", \"rows\": 15,\n
\"fields\": [\n    {\n        \"column\": \"Movie Title\", \n
\"properties\": {\n        \"dtype\": \"string\", \n
\"num_unique_values\": 15,\n        \"samples\": [\n
\"Spaceman\", \n        \"The Gray Man\", \n        \"Army of the\n
Dead\"\n        ],\n        \"semantic_type\": \"\", \n
\"description\": \"\"\n        }\n    },\n    {\n        \"column\":\n
\"Date\", \n        \"properties\": {\n        \"dtype\": \"string\", \n
\"num_unique_values\": 15,\n        \"samples\": [\n
\"2025-09-19\", \"2025-09-11\", \"2025-09-20\", \n
\"2025-09-20\", \"2025-09-07\", \"2025-09-26\", \n
\"2025-09-04\", \"2025-09-12\", \"2025-09-29\", \n
\"2025-09-01\", \n
\"2025-09-23\", \"2025-09-08\", \"2025-09-08\", \n
\"2025-09-10\", \"2025-09-13\", \"2025-09-13\", \n
\"2025-09-03\"\n        ],\n        \"semantic_type\": \"\", \n
\"description\": \"\"\n        }\n    },\n    {\n        \"column\": \"Region\", \n
\"properties\": {\n        \"dtype\": \"string\", \n
\"num_unique_values\": 15,\n        \"samples\": [\n
\"Australia\", \"USA\", \"USA\", \n
\"Japan\", \"Germany\", \"Australia\", \n
\"Canada\", \"Japan\", \"Brazil\", \n
\"India\", \n
\"Australia\", \"Japan\", \"UK\", \n
\"Japan\", \"India\", \"Germany\", \n
\"UK\"\n        ],\n        \"semantic_type\": \"\", \n
\"description\": \"\"\n        }\n    },\n    {\n        \"column\": \"Genre\", \n
\"properties\": {\n        \"dtype\": \"string\", \n
\"num_unique_values\": 15,\n        \"samples\": [\n
\"Horror\", \"Thriller\", \"Comedy\", \n
\"Horror\", \"Action\", \"Action\", \n
\"Thriller\", \"Horror\", \"Thriller\", \n
\"Romance\", \n
\"Drama\", \"Horror\", \"Action\", \n
\"Comedy\", \"Action\", \"Thriller\", \n
\"Action\"\n        ],\n        \"semantic_type\": \"\", \n
\"description\": \"\"\n        }\n    },\n    {\n        \"column\": \"Views (millions)\", \n
\"properties\": {\n        \"dtype\": \"number\", \n
\"std\": 21.364857660586996,\n        \"min\": 12.610000000000001,\n        \"max\": 84.84,\n        \"num_unique_values\": 15,\n        \"samples\": [\n
12.610000000000001,\n        53.15,\n
38.15\n        ],\n        \"semantic_type\": \"\", \n
\"description\": \"\"\n        }\n    },\n    {\n        \"column\":\n
\"Watch Time (hours)\", \n        \"properties\": {\n        \"dtype\":\n
\"number\", \n        \"std\": 56.05484383300749,\n        \"min\":\n

```

```

37.79,\n          \"max\": 238.15,\n          \"num_unique_values\": 15,\n\"samples\": [\n          37.79,\n          150.02,\n          98.44\n],\n          \"semantic_type\": \"\",\n          \"description\": \"\"\n}\n    },\n    {\n        \"column\": \"Revenue (USD millions)\",\n        \"properties\": {\n            \"dtype\": \"number\",\n            \"std\": 29.383107748371025,\n            \"min\": 20.1,\n            \"max\": 118.49,\n            \"num_unique_values\": 15,\n            \"samples\": [\n                20.1,\n                74.24000000000001,\n                49.09\n            ],\n            \"semantic_type\": \"\",\n            \"description\": \"\"\n        },\n        {\n            \"column\": \"Subscriber Type\",\n            \"properties\": {\n                \"dtype\": \"string\",\n                \"num_unique_values\": 15,\n                \"samples\": [\n                    \"BasicStandardStandard\",\n                    \"StandardStandardBasicBasicBasicStandard\",\n                    \"BasicPremiumPremiumPremiumPremiumStandardStandard\"\n                ],\n                \"semantic_type\": \"\",\n                \"description\": \"\"\n            }\n        }\n    }\n  ],\n  \"type\": \"dataframe\", \"variable_name\": \"grouped_df\"}

```

```

agg_df = df.groupby('Movie Title').agg({'Views (millions)': 'mean'})
agg_df

```

```

{"summary": "{\n  \"name\": \"agg_df\",\n  \"rows\": 15,\n  \"fields\": [\n    {\n      \"column\": \"Movie Title\",\n      \"properties\": {\n        \"dtype\": \"string\",\n        \"num_unique_values\": 15,\n        \"samples\": [\n          \"Spaceman\",\n          \"The Gray Man\",\n          \"Army of the Dead\"\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      },\n      {\n        \"column\": \"Views (millions)\",\n        \"properties\": {\n          \"dtype\": \"number\",\n          \"std\": 2.132778921632674,\n          \"min\": 3.68,\n          \"max\": 11.114,\n          \"num_unique_values\": 15,\n          \"samples\": [\n            4.2033333333333334,\n            7.5928571428571425,\n            5.45\n          ],\n          \"semantic_type\": \"\",\n          \"description\": \"\"\n        }\n      }\n    ]\n  },\n  \"type\": \"dataframe\", \"variable_name\": \"agg_df\"}

```

```

count_df = df.groupby('Movie Title').count()
count_df

```

```

{"summary": "{\n  \"name\": \"count_df\",\n  \"rows\": 15,\n  \"fields\": [\n    {\n      \"column\": \"Movie Title\",\n      \"properties\": {\n        \"dtype\": \"string\",\n        \"num_unique_values\": 15,\n        \"samples\": [\n          \"Spaceman\",\n          \"The Gray Man\",\n          \"Army of the Dead\"\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      },\n      {\n        \"column\": \"Date\",\n        \"properties\": {\n          \"dtype\": \"number\",\n          \"std\": 2,\n          \"min\": 3,\n          \"max\": 10,\n          \"num_unique_values\": 7,\n          \"samples\": [\n            7,\n            8,\n            9\n          ],\n          \"semantic_type\": \"\",\n          \"description\": \"\"\n        },\n        {\n          \"column\": \"Views (millions)\",\n          \"properties\": {\n            \"dtype\": \"number\",\n            \"std\": 2.132778921632674,\n            \"min\": 3.68,\n            \"max\": 11.114,\n            \"num_unique_values\": 15,\n            \"samples\": [\n              4.2033333333333334,\n              7.5928571428571425,\n              5.45\n            ],\n            \"semantic_type\": \"\",\n            \"description\": \"\"\n          }\n        }\n      }\n    ]\n  },\n  \"type\": \"dataframe\", \"variable_name\": \"count_df\"}

```



```

{"num_unique_values": 9, "samples": ["2025-09-11", "2025-09-01", "2025-09-15"], "semantic_type": "", "description": ""}, {"column": "Region", "properties": {"dtype": "category", "num_unique_values": 4, "samples": ["Brazil", "India", "Australia"], "semantic_type": "", "description": ""}, {"column": "Genre", "properties": {"dtype": "category", "num_unique_values": 3, "samples": ["Action", "Comedy", "Documentary"], "semantic_type": "", "description": ""}, {"column": "Views (millions)", "properties": {"dtype": "number", "std": 1.815361698185267, "min": 0.6, "max": 5.91, "num_unique_values": 15, "samples": [1.32, 0.79, 1.61], "semantic_type": "", "description": ""}, {"column": "Watch Time (hours)", "properties": {"dtype": "number", "std": 4.083989995790182, "min": 1.18, "max": 13.92, "num_unique_values": 15, "samples": [2.7, 1.77, 3.67], "semantic_type": "", "description": ""}, {"column": "Revenue (USD millions)", "properties": {"dtype": "number", "std": 2.034061618207931, "min": 0.92, "max": 7.52, "num_unique_values": 14, "samples": [2.29, 5.75, 2.41], "semantic_type": "", "description": ""}, {"column": "Subscriber Type", "properties": {"dtype": "category", "num_unique_values": 1, "samples": ["Basic"], "semantic_type": "", "description": ""}], "type": "dataframe", "variable_name": "min_df"}

```

```

df_cleaned = df.dropna()
df_cleaned

```

```

{"summary": {"name": "df_cleaned", "rows": 100, "fields": [{"column": "Date", "properties": {"dtype": "object", "num_unique_values": 28, "samples": ["2025-09-19", "2025-09-26", "2025-09-07"], "semantic_type": "", "description": ""}], "column": "Region", "properties": {"dtype": "category", "num_unique_values": 8, "samples": ["Australia", "USA", "India"], "semantic_type": "", "description": ""}, {"column": "Genre", "properties": {"dtype": "category", "num_unique_values": 3, "samples": ["Action", "Comedy", "Documentary"], "semantic_type": "", "description": ""}, {"column": "Views (millions)", "properties": {"dtype": "number", "std": 1.815361698185267, "min": 0.6, "max": 5.91, "num_unique_values": 15, "samples": [1.32, 0.79, 1.61], "semantic_type": "", "description": ""}, {"column": "Watch Time (hours)", "properties": {"dtype": "number", "std": 4.083989995790182, "min": 1.18, "max": 13.92, "num_unique_values": 15, "samples": [2.7, 1.77, 3.67], "semantic_type": "", "description": ""}, {"column": "Revenue (USD millions)", "properties": {"dtype": "number", "std": 2.034061618207931, "min": 0.92, "max": 7.52, "num_unique_values": 14, "samples": [2.29, 5.75, 2.41], "semantic_type": "", "description": ""}, {"column": "Subscriber Type", "properties": {"dtype": "category", "num_unique_values": 1, "samples": ["Basic"], "semantic_type": "", "description": ""}], "type": "dataframe", "variable_name": "min_df"}

```

```

\"column\": \"Movie Title\",
\"dtype\": \"category\",
\"num_unique_values\": 15,
\"samples\": [
  \"Triple Frontier\",
  \"Bird Box\",
  \"Don't Look Up\"
],
\"semantic_type\": \"\",
\"description\": \"\"
},
{
\"column\": \"Genre\",
\"dtype\": \"category\",
\"num_unique_values\": 7,
\"samples\": [
  \"Drama\",
  \"Thriller\",
  \"Documentary\"
],
\"semantic_type\": \"\",
\"description\": \"\"
},
{
\"column\": \"Views (millions)\",
\"dtype\": \"number\",
\"std\": 4.1969949266280295,
\"min\": 0.6,
\"max\": 15.0,
\"num_unique_values\": 98,
\"samples\": [
  8.0,
  6.64,
  10.22
],
\"semantic_type\": \"\",
\"description\": \"\"
},
{
\"column\": \"Watch Time (hours)\",
\"dtype\": \"number\",
\"std\": 12.333858748431595,
\"min\": 1.18,
\"max\": 49.83,
\"num_unique_values\": 97,
\"samples\": [
  14.76,
  5.94,
  23.83
],
\"semantic_type\": \"\",
\"description\": \"\"
},
{
\"column\": \"Revenue (USD millions)\",
\"dtype\": \"number\",
\"std\": 6.458923714168993,
\"min\": 0.92,
\"max\": 27.8,
\"num_unique_values\": 93,
\"samples\": [
  2.82,
  5.59,
  8.54
],
\"semantic_type\": \"\",
\"description\": \"\"
},
{
\"column\": \"Subscriber Type\",
\"dtype\": \"category\",
\"num_unique_values\": 3,
\"samples\": [
  \"Standard\",
  \"Premium\",
  \"Basic\"
],
\"semantic_type\": \"\",
\"description\": \"\"
}
],
\"type\": \"dataframe\",
\"variable_name\": \"df_cleaned\"

```

```

df_filled = df.fillna(0)
df_filled

```

```

{
  \"summary\": {
    \"name\": \"df_filled\",
    \"rows\": 100,
    \"fields\": [
      {
        \"column\": \"Date\",
        \"dtype\": \"object\",
        \"num_unique_values\": 28,
        \"samples\": [
          \"2025-09-19\",
          \"2025-09-26\",
          \"2025-09-07\"
        ],
        \"semantic_type\": \"\",
        \"description\": \"\"
      }
    ],
    \"column\": \"Region\",
    \"dtype\": \"category\",
    \"num_unique_values\": 8,
    \"samples\": [
      \"Australia\",
      \"USA\",
      \"India\"
    ],
    \"semantic_type\": \"\",
    \"description\": \"\"
  },
  \"column\": \"Movie Title\",
  \"dtype\": \"category\",
  \"num_unique_values\": 15,

```

```

\"samples\": [\n          \"Triple Frontier\", \n          \"Bird Box\", \n          \"Don't Look Up\" ], \n
\"semantic_type\": \"\", \n
\"description\": \"\" \n
} \n
{ \n
\"column\": \"Genre\", \n
\"properties\": { \n
\"dtype\": \"category\", \n
\"num_unique_values\": 7, \n
\"samples\": [\n          \"Drama\", \n          \"Thriller\", \n          \"Documentary\" ], \n
\"semantic_type\": \"\", \n
\"description\": \"\" \n
} \n
}, \n
{ \n
\"column\": \"Views (millions)\", \n
\"properties\": { \n
\"dtype\": \"number\", \n
\"std\": 4.1969949266280295, \n
\"min\": 0.6, \n
\"max\": 15.0, \n
\"num_unique_values\": 98, \n
\"samples\": [\n          8.0, \n          6.64, \n          10.22 \n
], \n
\"semantic_type\": \"\", \n
\"description\": \"\" \n
} \n
}, \n
{ \n
\"column\": \"Watch Time (hours)\", \n
\"properties\": { \n
\"dtype\": \"number\", \n
\"std\": 12.333858748431595, \n
\"min\": 1.18, \n
\"max\": 49.83, \n
\"num_unique_values\": 97, \n
\"samples\": [\n          14.76, \n          5.94, \n          23.83 \n
], \n
\"semantic_type\": \"\", \n
\"description\": \"\" \n
} \n
}, \n
{ \n
\"column\": \"Revenue (USD millions)\", \n
\"properties\": { \n
\"dtype\": \"number\", \n
\"std\": 6.458923714168993, \n
\"min\": 0.92, \n
\"max\": 27.8, \n
\"num_unique_values\": 93, \n
\"samples\": [\n          2.82, \n          5.59, \n          8.54 \n
], \n
\"semantic_type\": \"\", \n
\"description\": \"\" \n
} \n
}, \n
{ \n
\"column\": \"Subscriber Type\", \n
\"properties\": { \n
\"dtype\": \"category\", \n
\"num_unique_values\": 3, \n
\"samples\": [\n          \"Standard\", \n          \"Premium\", \n          \"Basic\" ], \n
\"semantic_type\": \"\", \n
\"description\": \"\" \n
} \n
} ] \n
n} \", \"type\": \"dataframe\", \"variable_name\": \"df_filled\"}

```

```

df_replaced = df.replace({'NaN': 0})
df_replaced

```

```

{ \"summary\": { \n
\"name\": \"df_replaced\", \n
\"rows\": 100, \n
\"fields\": [\n          { \n
\"column\": \"Date\", \n
\"properties\": { \n
\"dtype\": \"object\", \n
\"num_unique_values\": 28, \n
\"samples\": [\n          \"2025-09-19\", \n          \"2025-09-26\", \n          \"2025-09-07\" \n
], \n
\"semantic_type\": \"\", \n
\"description\": \"\" \n
} \n
}, \n
{ \n
\"column\": \"Region\", \n
\"properties\": { \n
\"dtype\": \"category\", \n
\"num_unique_values\": 8, \n
\"samples\": [\n          \"Australia\", \n          \"USA\", \n          \"India\" \n
], \n
\"semantic_type\": \"\", \n
\"description\": \"\" \n
} \n
}, \n
{ \n
\"column\": \"Movie Title\", \n
\"properties\": { \n
\"dtype\": \"category\", \n
\"num_unique_values\": 15, \n
\"samples\": [\n          \"Triple Frontier\", \n          \"Bird Box\", \n          \"Don't Look Up\" \n
], \n

```

```

{"semantic_type": "\\",
  "description": "\\",
  "column": "Genre",
  "properties": {
    "dtype": "category",
    "num_unique_values": 7,
    "samples": [
      "Drama",
      "Thriller",
      "Documentary"
    ],
    "semantic_type": "\\",
    "description": "\\",
    "column": "Views (millions)",
    "properties": {
      "dtype": "number",
      "std": 4.1969949266280295,
      "min": 0.6,
      "max": 15.0,
      "num_unique_values": 98,
      "samples": [
        8.0,
        6.64,
        10.22
      ]
    },
    "semantic_type": "\\",
    "description": "\\",
    "column": "Watch Time (hours)",
    "properties": {
      "dtype": "number",
      "std": 12.333858748431595,
      "min": 1.18,
      "max": 49.83,
      "num_unique_values": 97,
      "samples": [
        14.76,
        5.94,
        23.83
      ]
    },
    "semantic_type": "\\",
    "description": "\\",
    "column": "Revenue (USD millions)",
    "properties": {
      "dtype": "number",
      "std": 6.458923714168993,
      "min": 0.92,
      "max": 27.8,
      "num_unique_values": 93,
      "samples": [
        2.82,
        5.59,
        8.54
      ]
    },
    "semantic_type": "\\",
    "description": "\\",
    "column": "Subscriber Type",
    "properties": {
      "dtype": "category",
      "num_unique_values": 3,
      "samples": [
        "Standard",
        "Premium",
        "Basic"
      ]
    },
    "semantic_type": "\\",
    "description": "\\",
    "column": " "
  }
},
{"type": "dataframe", "variable_name": "df_replaced"}

```

```

df['Contains_a'] = df['Movie Title'].str.contains('D')
df

```

```

{"summary": {
  "name": "df",
  "rows": 100,
  "fields": [
    {
      "column": "Date",
      "properties": {
        "dtype": "object",
        "num_unique_values": 28,
        "samples": [
          "2025-09-19",
          "2025-09-26",
          "2025-09-07"
        ],
        "semantic_type": "\\",
        "description": "\\",
        "column": "Region",
        "properties": {
          "dtype": "category",
          "num_unique_values": 8,
          "samples": [
            "Australia",
            "USA",
            "India"
          ],
          "semantic_type": "\\",
          "description": "\\",
          "column": "Movie Title",
          "properties": {
            "dtype": "category",
            "num_unique_values": 15,
            "samples": [
              "Triple Frontier",
              "Bird Box",
              "Don't Look Up"
            ],
            "semantic_type": "\\",
            "description": "\\",
            "column": "Genre",
            "properties": {
              "dtype": "category",
              "num_unique_values": 

```

```

7,\n          \"samples\": [\n          \"Drama\", \n          \"Thriller\", \n          \"Documentary\" ], \n          \"semantic_type\": \"\", \n          \"description\": \"\" \n      }, \n      {\n          \"column\": \"Views (millions)\", \n          \"properties\": {\n              \"dtype\": \"number\", \n              \"std\": 4.1969949266280295, \n              \"min\": 0.6, \n              \"max\": 15.0, \n              \"num_unique_values\": 98, \n              \"samples\": [\n                  8.0, \n                  6.64, \n                  10.22 \n              ], \n              \"semantic_type\": \"\", \n              \"description\": \"\" \n          }, \n          {\n              \"column\": \"Watch Time (hours)\", \n              \"properties\": {\n                  \"dtype\": \"number\", \n                  \"std\": 12.333858748431595, \n                  \"min\": 1.18, \n                  \"max\": 49.83, \n                  \"num_unique_values\": 97, \n                  \"samples\": [\n                      14.76, \n                      5.94, \n                      23.83 \n                  ], \n                  \"semantic_type\": \"\", \n                  \"description\": \"\" \n              }, \n              {\n                  \"column\": \"Revenue (USD millions)\", \n                  \"properties\": {\n                      \"dtype\": \"number\", \n                      \"std\": 6.458923714168993, \n                      \"min\": 0.92, \n                      \"max\": 27.8, \n                      \"num_unique_values\": 93, \n                      \"samples\": [\n                          2.82, \n                          5.59, \n                          8.54 \n                      ], \n                      \"semantic_type\": \"\", \n                      \"description\": \"\" \n                  }, \n                  {\n                      \"column\": \"Subscriber Type\", \n                      \"properties\": {\n                          \"dtype\": \"category\", \n                          \"num_unique_values\": 3, \n                          \"samples\": [\n                              \"Standard\", \n                              \"Premium\", \n                              \"Basic\" \n                          ], \n                          \"semantic_type\": \"\", \n                          \"description\": \"\" \n                      }, \n                      {\n                          \"column\": \"Contains_a\", \n                          \"properties\": {\n                              \"dtype\": \"boolean\", \n                              \"num_unique_values\": 2, \n                              \"samples\": [\n                                  false, \n                                  true \n                              ], \n                              \"semantic_type\": \"\", \n                              \"description\": \"\" \n                          }, \n                          {\n                              \"column\": \"Contains_Project\", \n                              \"properties\": {\n                                  \"dtype\": \"boolean\", \n                                  \"num_unique_values\": 2, \n                                  \"samples\": [\n                                      false, \n                                      true \n                                  ], \n                                  \"semantic_type\": \"\", \n                                  \"description\": \"\" \n                              } \n                          } \n                      } \n                  } \n              } \n          } \n      ], \n      {\n          \"column\": \"Contains_Project\", \n          \"properties\": {\n              \"dtype\": \"boolean\", \n              \"num_unique_values\": 2, \n              \"samples\": [\n                  false, \n                  true \n              ], \n              \"semantic_type\": \"\", \n              \"description\": \"\" \n          } \n      } \n  ], \n  \"type\": \"dataframe\", \"variable_name\": \"df\"}

```

```

counts = df['Movie Title'].value_counts( )
counts

```

```

Movie Title
The Old Guard    10
Red Notice       9
The Adam Project  8
Bird Box          8
Enola Holmes 2   8
Triple Frontier   8
Heart of Stone    8
The Gray Man      7
Project Power     7
Army of the Dead  7

```

```

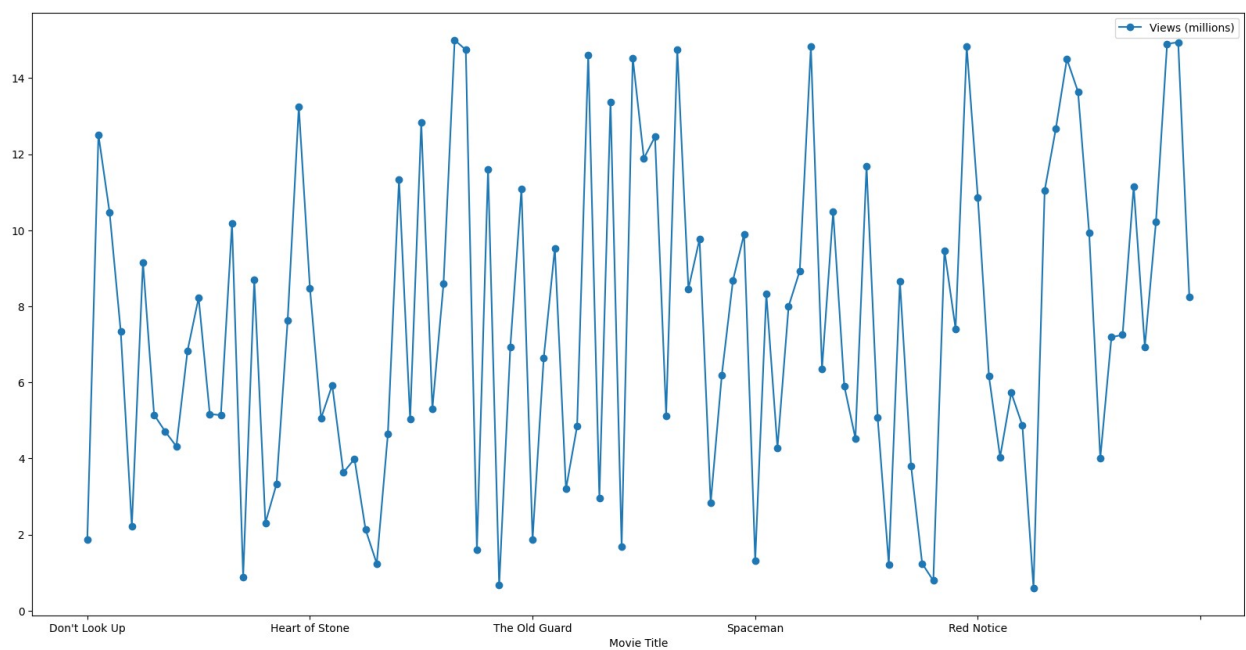
Don't Look Up        6
Extraction 2         5
Glass Onion          3
The Mother           3
Spaceman             3
Name: count, dtype: int64

```

```
import matplotlib.pyplot as plt
```

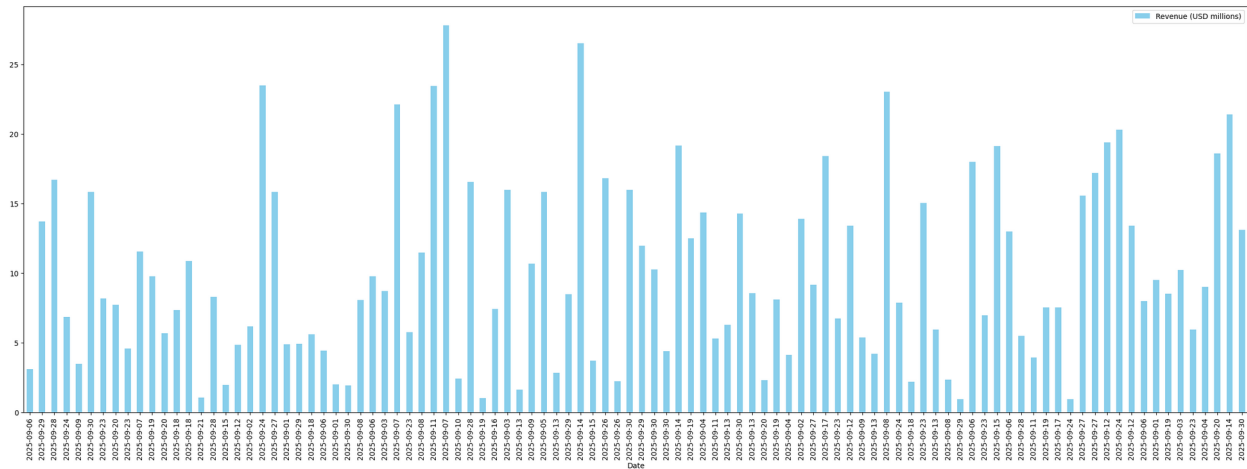
```
df.plot(x='Movie Title', y='Views (millions)', kind='line',
marker='o',figsize=(20,10))
```

```
<Axes: xlabel='Movie Title'>
```



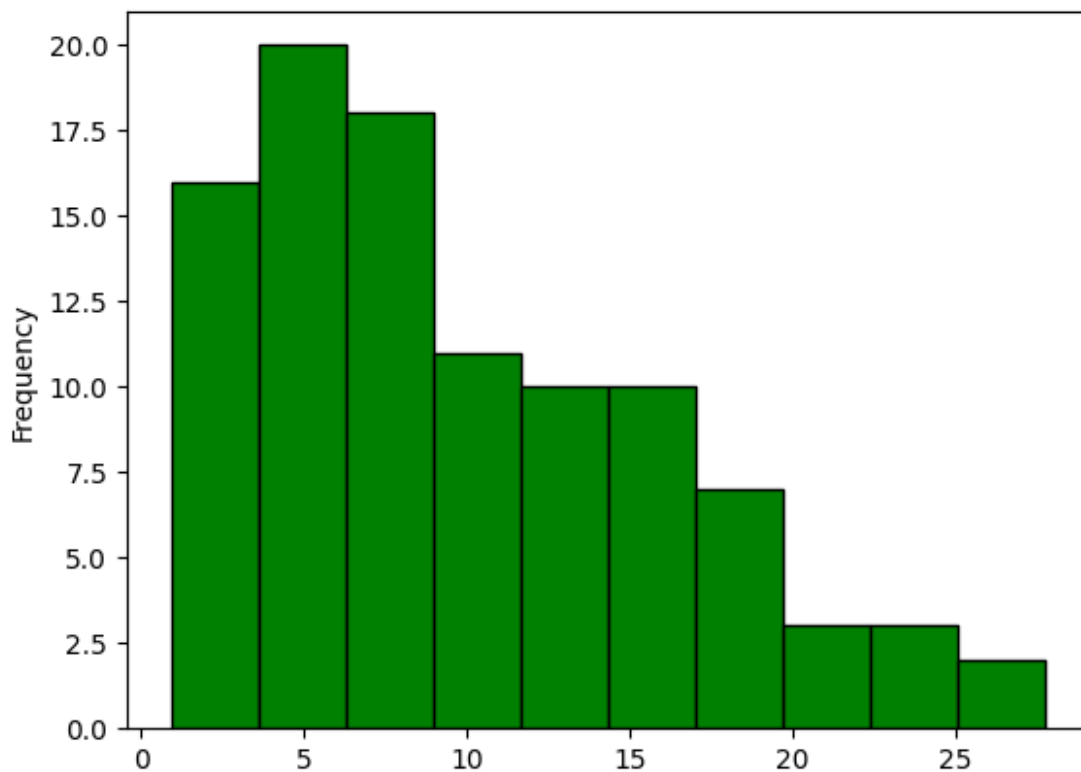
```
df.plot(x='Date', y='Revenue (USD millions)', kind='bar',
color='skyblue',figsize=(30,10))
```

```
<Axes: xlabel='Date'>
```



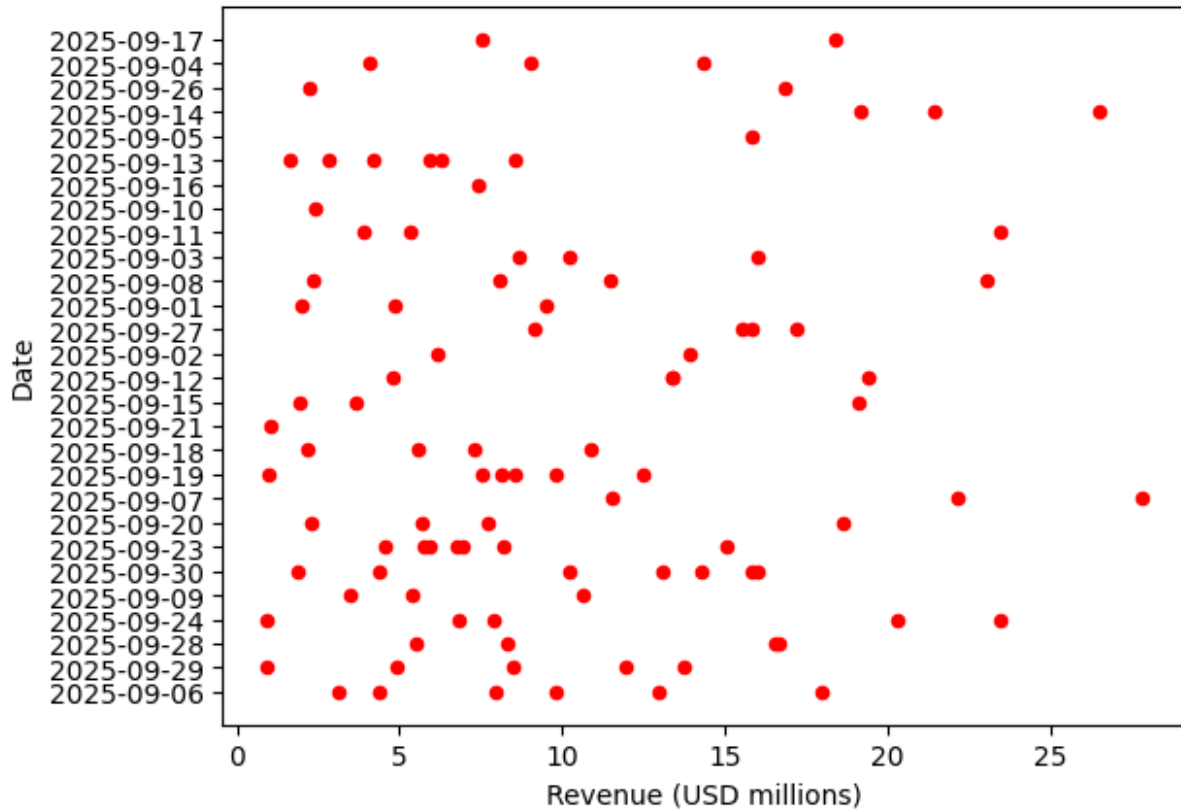
```
df['Revenue (USD millions)'].plot(kind='hist',color='green',
edgecolor='black')
```

```
<Axes: ylabel='Frequency'>
```

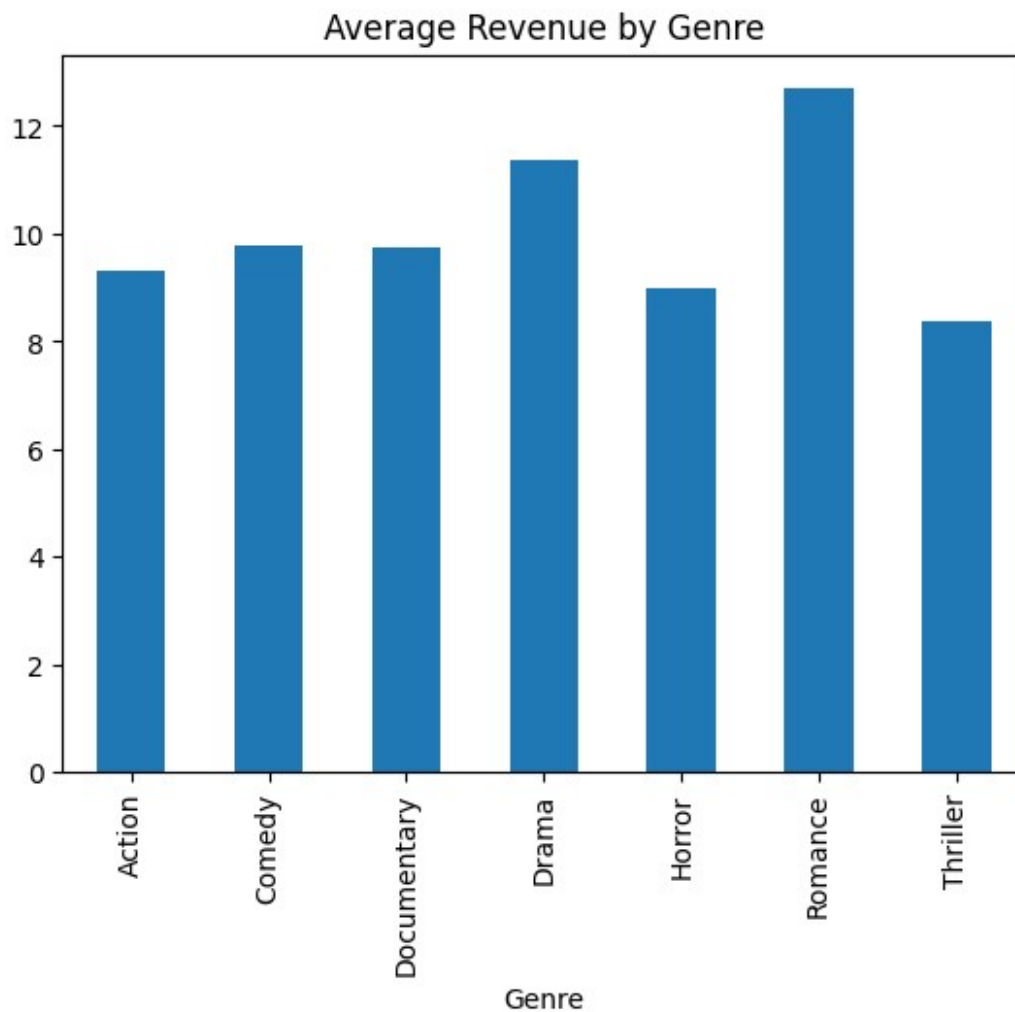


```
df.plot(x='Revenue (USD millions)', y='Date', kind='scatter',
color='red')
```

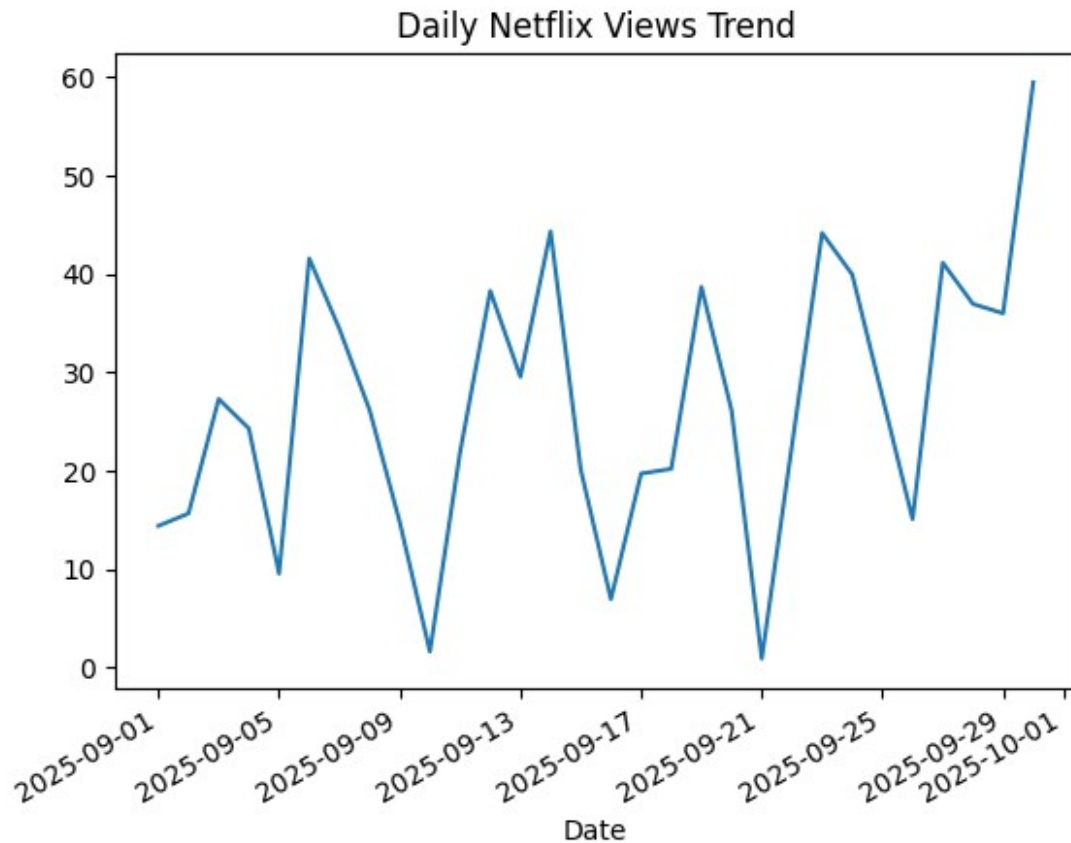
```
<Axes: xlabel='Revenue (USD millions)', ylabel='Date'>
```



```
avg_revenue = df.groupby('Genre')['Revenue (USD millions)'].mean()
avg_revenue.plot(kind='bar', title='Average Revenue by Genre')
<Axes: title={'center': 'Average Revenue by Genre'}, xlabel='Genre'>
```

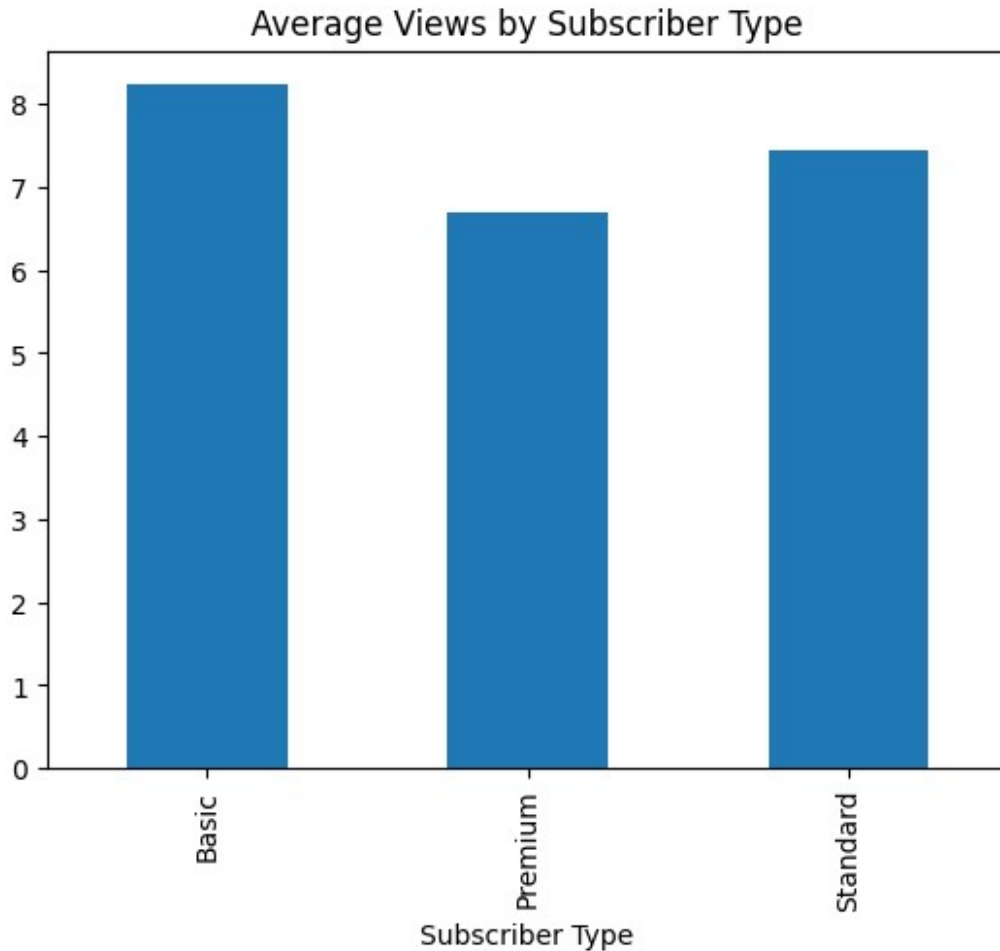


```
df['Date'] = pd.to_datetime(df['Date'])
df.groupby('Date')['Views (millions)'].sum().plot(kind='line',
title='Daily Netflix Views Trend')
<Axes: title={'center': 'Daily Netflix Views Trend'}, xlabel='Date'>
```



```
df.groupby('Subscriber Type')['Views  
(millions)'].mean().plot(kind='bar', title='Average Views by  
Subscriber Type')
```

```
<Axes: title={'center': 'Average Views by Subscriber Type'},  
xlabel='Subscriber Type'>
```



#Which genre has the highest total views across all regions?

```
df.groupby('Genre')['Views  
(millions)'].sum().sort_values(ascending=False).head(1)
```

Genre

Action 147.91

Name: Views (millions), dtype: float64

#Find the correlation between views, watch time, and revenue.

```
df[['Views (millions)', 'Watch Time (hours)', 'Revenue (USD  
millions)']].corr()
```

```
{"summary": "{\n  \"name\": \"df[['Views (millions)', 'Watch Time  
(hours)', 'Revenue (USD millions)']\"\n\", \"rows\": 3,\n  \"fields\": [\n    {\n      \"column\": \"Views (millions)\"\n    },\n    {\n      \"column\": \"Watch Time (hours)\"\n    },\n    {\n      \"column\": \"Revenue (USD millions)\"\n    }\n  ],\n  \"properties\": {\n    \"Views (millions)\": {\n      \"dtype\": \"number\",\n      \"std\": 0.0539292619312695,\n      \"min\": 0.9048633112380353,\n      \"max\": 1.0,\n      \"num_unique_values\": 3,\n      \"samples\": [\n        1.0,\n        0.9048633112380353,\n        0.9084219553711039\n      ],\n      \"semantic_type\": \"\",\n      \"description\": \"\"\n    },\n    \"Watch Time (hours)\": {\n      \"dtype\": \"number\",\n      \"std\": 0.0539292619312695,\n      \"min\": 0.9048633112380353,\n      \"max\": 1.0,\n      \"num_unique_values\": 3,\n      \"samples\": [\n        1.0,\n        0.9048633112380353,\n        0.9084219553711039\n      ],\n      \"semantic_type\": \"\",\n      \"description\": \"\"\n    },\n    \"Revenue (USD millions)\": {\n      \"dtype\": \"number\",\n      \"std\": 0.0539292619312695,\n      \"min\": 0.9048633112380353,\n      \"max\": 1.0,\n      \"num_unique_values\": 3,\n      \"samples\": [\n        1.0,\n        0.9048633112380353,\n        0.9084219553711039\n      ],\n      \"semantic_type\": \"\",\n      \"description\": \"\"\n    }\n  }\n}
```

```

{"Watch Time (hours)": 0.811425264849265, "properties": {"dtype": "number", "std": 0.094288642651798, "min": 0.811425264849265, "max": 1.0, "num_unique_values": 3, "samples": [0.9048633112380353, 1.0, 0.811425264849265]}, "semantic_type": "", "description": ""}, {"column": "Revenue (USD millions)", "properties": {"dtype": "number", "std": 0.09430034196421852, "min": 0.811425264849265, "max": 1.0, "num_unique_values": 3, "samples": [0.9084219553711039, 0.811425264849265, 1.0]}, "semantic_type": "", "description": ""}]}, {"type": "dataframe"}

```

#Identify the most profitable day (highest total revenue).

```

df['Date'] = pd.to_datetime(df['Date'])
df.groupby('Date')['Revenue (USD
millions)'].sum().sort_values(ascending=False).head(1)

```

```

Date
2025-09-30    75.71
Name: Revenue (USD millions), dtype: float64

```

#Create a scatter plot between Watch Time and Revenue.

```

plt.scatter(df['Watch Time (hours)'], df['Revenue (USD millions)'],
alpha=0.7)
plt.title('Watch Time vs Revenue')
plt.xlabel('Watch Time (hours)')
plt.ylabel('Revenue (USD millions)')
plt.show()

```



```
#Determine which region has the most balanced genre performance (least variance in revenue).
```

```
df.groupby('Region')['Revenue (USD  
millions)'].var().sort_values().head(1)
```

```
Region
```

```
Brazil    28.907364
```

```
Name: Revenue (USD millions), dtype: float64
```

```
#Plot total Netflix views by region (bar graph).
```

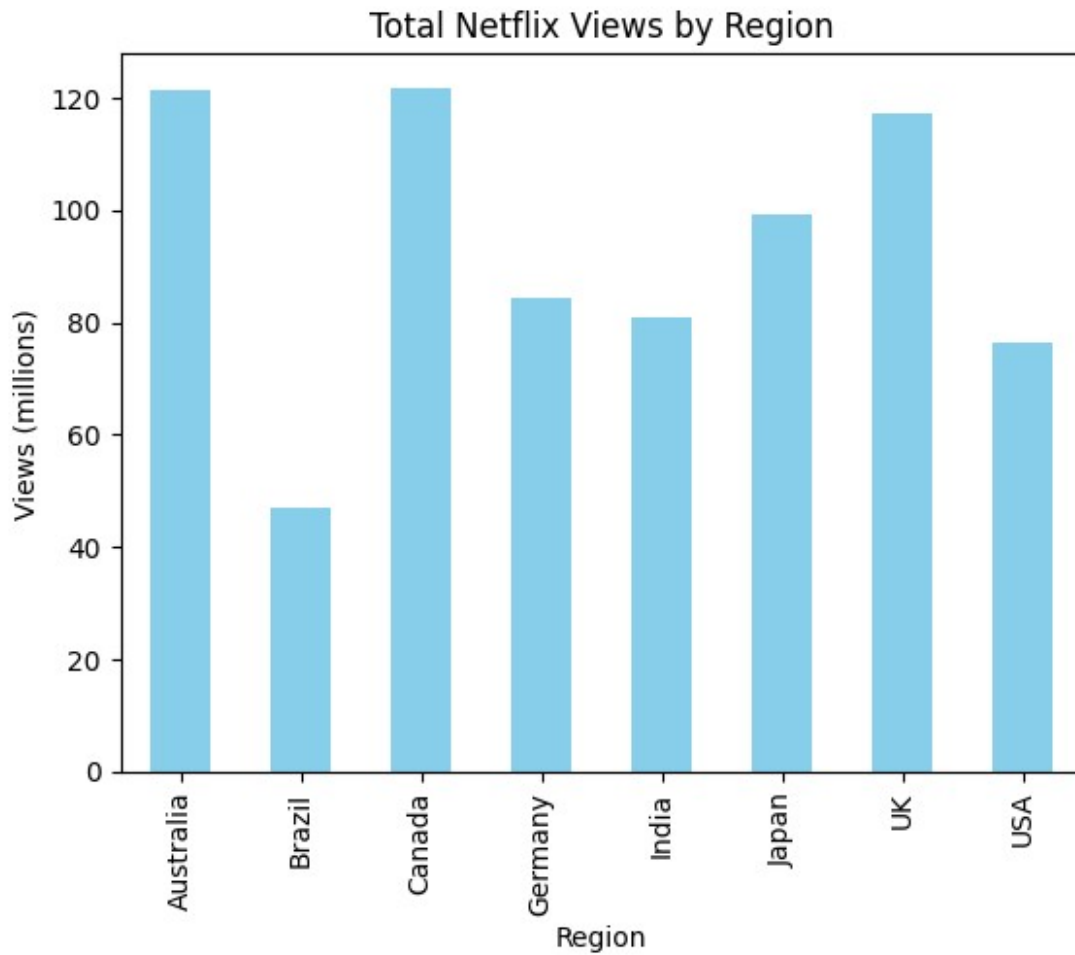
```
df.groupby('Region')['Views (millions)'].sum().plot(kind='bar',  
color='skyblue')
```

```
plt.title('Total Netflix Views by Region')
```

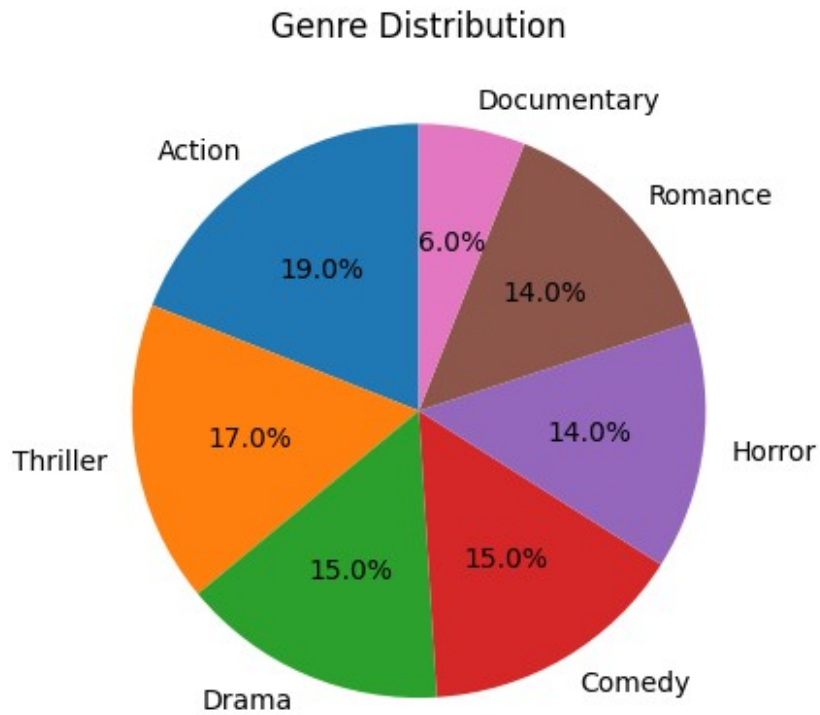
```
plt.ylabel('Views (millions)')
```

```
plt.xlabel('Region')
```

```
plt.show()
```



```
#Pie chart showing percentage share of genres.  
df['Genre'].value_counts().plot(kind='pie', autopct='%1.1f%%',  
startangle=90)  
plt.title('Genre Distribution')  
plt.ylabel('')  
plt.show()
```



```
#Multi-bar chart: revenue per genre by subscriber type.  
df.groupby(['Genre', 'Subscriber Type'])['Revenue (USD  
millions)'].mean().unstack().plot(kind='bar')  
plt.title('Revenue per Genre by Subscriber Type')  
plt.ylabel('Revenue (USD millions)')  
plt.show()
```

