DATE: / /

Name : Danijal Saeed
Sap id: 53937
Section: BS Data Science
Course: Advance statistics

Assignment

Pandas Python Libraries.

Importing Pandas.

```
import pandas as pd.
```

pd.Series.

```
s = pd.Series([10, 20, 30])
```

pd.DataFrame().

```
df = pd.DataFrame({"A": [1, 2], "B": [3, 4]})
```

pd.read_csv().

```
df = pd.read_csv("data.csv").
```

pd.read_excel().

```
df = pd.read_excel("file.xlsx")
```

DATE: / /

df.to_csv().

```
df.to_csv("output.csv", index=False)
```

df.to_excel().

```
df.to_excel("output.xlsx", index=False)
```

df.head()

```
df.head(5)
```

df.tail()

```
df.tail(3)
```

df.info().

```
df.info()
```

df.describe()

```
df.describe()
```

df.shape

```
df.shape
```

DATE: / /

df.colums

df.colums

df.dtyes

df.dtypes

df['col']

df['Age']

df.[['col1','col2']]

df[['Name', Age']]

df.loc[]

df.loc[0:5,['Name', Age']]

df.iloc[]

df.iloc[0:3, 1:4]

df[df['Age']>30]

df[df['Age']>30]

DATE: / /

df.assign().

```
df = df.assign(Total = df['A'] + df['B'])
```

df.query()

```
df.query("Age > 30 and City == 'Lahore'")
```

df.rename()

```
df.rename(columns={'A': 'Age'}, inplace=True)
```

df.drop()

```
df.drop(columns=['City'], inplace=True)
```

df.replace()

```
df['City'].replace("Lahore", "LA", inplace=True)
```

df.fillna()

```
df.fillna(0)
```

df.dropna()

```
df.dropna()
```

DATE / /

df.sort_values()

```
df.sort_values("Age", ascending=False)
```

df.groupby()

```
df.groupby("City")["Salary"].mean()
```

df.agg()

```
df.agg({"Age":"mean", "Salary":"sum"})
```

pd.concat()

```
pd.concat([df1, df2], axis=0)
```

pd.merge()

```
pd.merge(df1, df2, on='ID',
         how='inner')
```

df.join()

```
df.join(df2, lsuffix='-A', rsuffix=
        'B')
```

pd.to_datetime()

```
df['Date'] = pd.to_datetime
             (df['Date'])
```

DATE: / /

df.set_index()

```
df.set_index('Date', inplace=True)
```

df.apply()

```
df['Age'] = df['Age'].apply(lambda x: x*2)
```

df.value_counts()

```
df['City'].value_counts()
```

df.unique()

```
df.unique()
```

df.sample()

```
df.sample(3)
```

df.corr()

```
df.corr()
```

df.pivot_table()

```
df.pivot_table(values='sales',
               index='City', aggfunc='sum')
```