# NATIONAL UNIVERSITY OF MODERN LANGUAGES ISLAMABAD



## Data Mining (LAB)

### Lab Task: 08

**Submitted to**
Dr. Moiz-ullah Ghouri

**Submitted By**
Junaid Asif
(BSAI-144)

**Submission Date:** December 10, 2024

```python
import numpy as np
import pandas as pd
```

```python
dates = pd.read_csv('dates.csv')
```

```python
dates
```

|   | month_day_year | day_month_year | date_time | year_month_day |
|---|---|---|---|---|
| 0 | 04/22/96 | 22-Apr-96 | Tue Aug 11 09:50:35 1996 | 2007-06-22 |
| 1 | 04/23/96 | 23-Apr-96 | Tue May 12 19:50:35 2016 | 2017-01-09 |
| 2 | 05/14/96 | 14-May-96 | Mon Oct 14 09:50:35 2017 | 1998-04-12 |
| 3 | 05/15/96 | 15-May-96 | Tue Jan 11 09:50:35 2018 | 2027-07-22 |
| 4 | 05/16/01 | 16-May-01 | Fri Mar 11 07:30:36 2019 | 1945-11-15 |
| 5 | 05/17/02 | 17-May-02 | Tue Aug 11 09:50:35 2020 | 1942-06-22 |
| 6 | 05/18/03 | 18-May-03 | Wed Dec 21 09:50:35 2021 | 1887-06-13 |
| 7 | 05/19/04 | 19-May-04 | Tue Jan 11 09:50:35 2022 | 1912-01-25 |
| 8 | 05/20/05 | 20-May-05 | Sun Jul 10 19:40:25 2023 | 2007-06-22 |

```python
for col in dates:
    print(type(dates[col][3]))
```

```
<class 'pandas._libs.tslibs.timestamps.Timestamp'>
<class 'pandas._libs.tslibs.timestamps.Timestamp'>
<class 'pandas._libs.tslibs.timestamps.Timestamp'>
<class 'pandas._libs.tslibs.timestamps.Timestamp'>
```

```python
dates = pd.read_csv('dates.csv', parse_dates=[0,1,2,3])
```

```
C:\Users\junai\AppData\Local\Temp\ipykernel_10012\1239452800.py:1: UserWarning: Could not infer
  dates = pd.read_csv('dates.csv', parse_dates=[0,1,2,3])
C:\Users\junai\AppData\Local\Temp\ipykernel_10012\1239452800.py:1: UserWarning: Could not infer
  dates = pd.read_csv('dates.csv', parse_dates=[0,1,2,3])
```

```python
dates['date_time'] = pd.to_datetime(dates['date_time'], format='%a %b %d %H:%M:%S %Y')
```

```python
column_1 = dates.iloc[:, 0]
pd.DataFrame({'year': column_1.dt.year,
              'month': column_1.dt.month,
              'day': column_1.dt.day,
              'hour': column_1.dt.hour,
              'dayofyear': column_1.dt.dayofyear,
              #'week': column_1.dt.isocalendar().week,
              'weekofyear': column_1.dt.isocalendar().week,
              #'dayofweek': column_1.dt.dayofweek,
              'weekday': column_1.dt.weekday,
              'quarter': column_1.dt.quarter,

})
```

| | year | month | day | hour | dayofyear | weekofyear | weekday | quarter |
|---|------|-------|-----|------|-----------|------------|---------|---------|
| 0 | 1996 | 4 | 22 | 0 | 113 | 17 | 0 | 2 |
| 1 | 1996 | 4 | 23 | 0 | 114 | 17 | 1 | 2 |
| 2 | 1996 | 5 | 14 | 0 | 135 | 20 | 1 | 2 |
| 3 | 1996 | 5 | 15 | 0 | 136 | 20 | 2 | 2 |
| 4 | 2001 | 5 | 16 | 0 | 136 | 20 | 2 | 2 |
| 5 | 2002 | 5 | 17 | 0 | 137 | 20 | 4 | 2 |
| 6 | 2003 | 5 | 18 | 0 | 138 | 20 | 6 | 2 |
| 7 | 2004 | 5 | 19 | 0 | 140 | 21 | 2 | 2 |
| 8 | 2005 | 5 | 20 | 0 | 140 | 20 | 4 | 2 |

```python
print(dates.iloc[1,0])
print(dates.iloc[3,0])
print(dates.iloc[3,0]-dates.iloc[1,0])
```

```
1996-04-23 00:00:00
1996-05-15 00:00:00
22 days 00:00:00
```