Urban Data Analysis

Week 12: Time Series Analysis & Forecasting

Today's Agenda

- INTRODUCTION TO TIME SERIES DATA

- KEY COMPONENTS OF TIME SERIES

- FORECASTING WITH STATSMODELS

- MODEL EVALUATION METRICS

- URBAN MOBILITY FORECASTING CASE STUDY

What is Time Series Analysis?

TIME SERIES DATA: DATA POINTS COLLECTED OR RECORDED AT SPECIFIC TIME INTERVALS

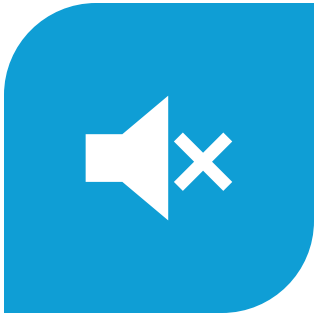
EXAMPLES: SUBWAY RIDERSHIP, DAILY AIR QUALITY

GOAL: UNDERSTAND PATTERNS AND FORECAST FUTURE VALUES

Time Series Components

1. TREND – LONG-TERM MOVEMENT

2. SEASONALITY – REPEATING CYCLES



3. NOISE – RANDOM FLUCTUATIONS

Urban Use Case Example

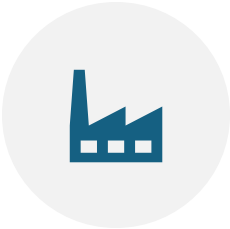
EXAMPLE:

CAN WE FORECAST NEXT WEEK'S SUBWAY RIDERSHIP USING PAST DAILY DATA?

USE GTFS OR MTA TURNSTILE DATA

Step 1: Import and Plot

```
import pandas as pd
```



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

```
ridership =
pd.read_csv("daily_subway_ridership.csv"
, parse_dates=['date'], index_col='date')
```

```
ridership['entries'].plot(figsize=(12,5),
title='Daily Subway Entries')
```

Step 2: Decompose Time Series

```
from statsmodels.tsa.seasonal import seasonal_decompose
```

```
result = seasonal_decompose(ridership['entries'], model='additive', period=7)
```

```
result.plot()
```

Step 3: Forecast with ARIMA

```
from statsmodels.tsa.arima.model import ARIMA
```

```
model = ARIMA(ridership['entries'], order=(1,1,1))
```

```
fit = model.fit()
```

```
forecast = fit.forecast(steps=7)
```

Step 4: Plot Forecast

plt.plot(ridership.index[-30:], ridership['entries'][-30:], label='Actual')

plt.plot(pd.date_range(start=ridership.index[-1], periods=8, freq='D')[1:], forecast, label='Forecast')

Step 5: Evaluation Metrics

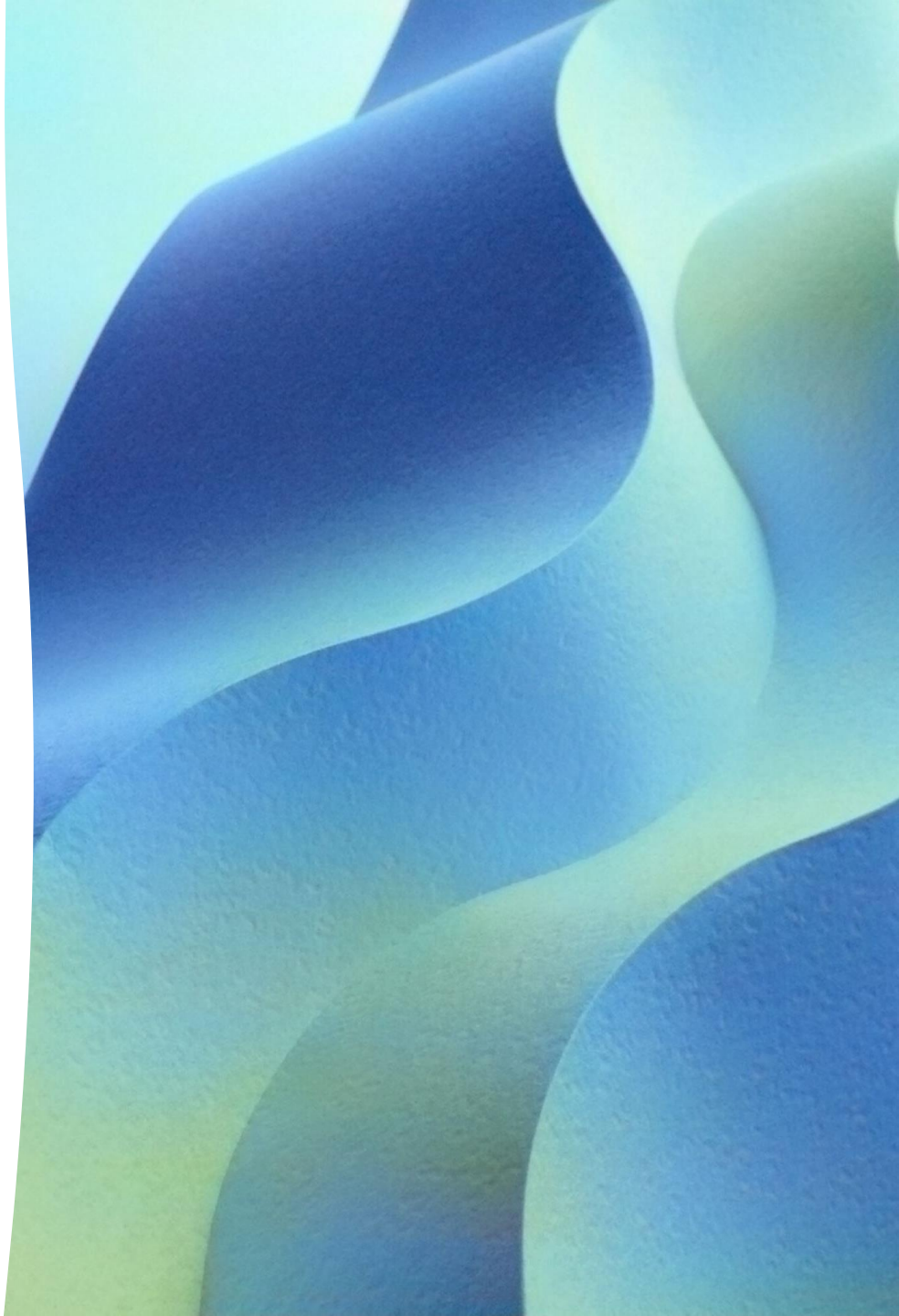
MAE, RMSE, MAPE

```
from sklearn.metrics import
mean_absolute_error, mean_squared_error
```

```
import numpy as np
```

```
mae = mean_absolute_error(actual,
forecast)
```

```
rmse = np.sqrt(mean_squared_error(actual,
forecast))
```

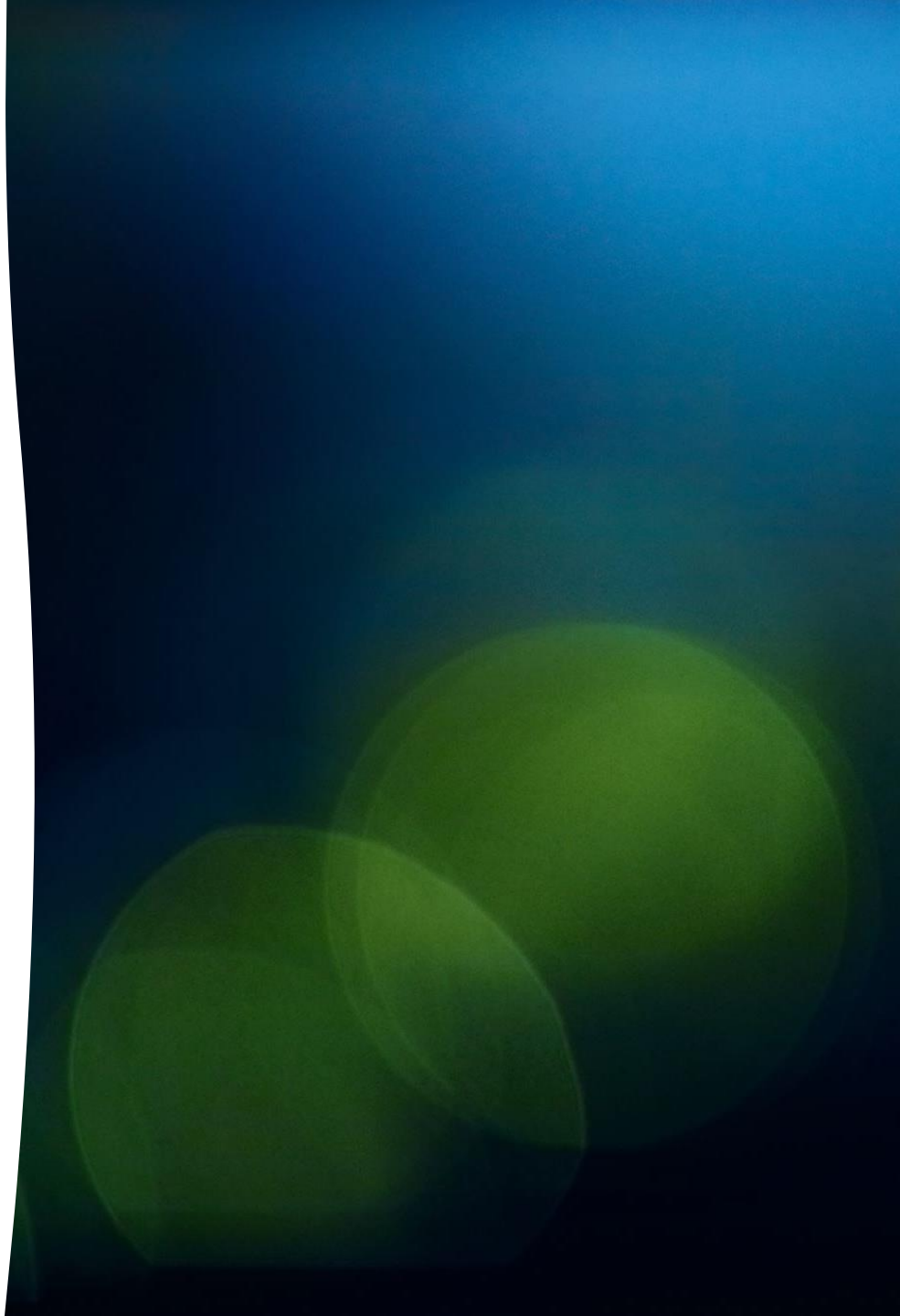


Challenges in Forecasting

- Data gaps (weekends/holidays)

- Anomalies (weather, construction)

- High variability in demand



Wrap-Up

- TIME SERIES ANALYSIS IS CRITICAL IN URBAN PLANNING

- HELPS FORECAST DEMAND AND SUPPORT DECISIONS

- EVALUATE ACCURACY CAREFULLY

Discussion Questions

- What time frequency is best: daily, weekly, monthly?

- How do shocks (e.g., COVID) affect results?

- How can forecasts be used by agencies?

Thank You!