



# Recitation 3

Data Analytics (18-787)

Friday 17 Feb, 2023



## Note on recitation slides

- Recitation slides are intended to be a guide on how to approach the assignment and not a prescription of exactly what to do
- There could be many approaches to any problem
- Seek to understand the problem and solve it instead of just trying to reproduce the steps listed in the slides
- To avoid overdependence on the slides, start your assignment early!
- If you have to choose between a “creative” approach, or following assignment instructions in the PDF, choose to follow PDF instructions **always**



# Assignment Objectives

- Explore and visualize time series data
- Investigate variability in time series data over different time periods
- Perform statistical hypothesis test
- Evaluate model performance based on benchmarking



## Question 1 : Time series analysis

- Download intraday 15-minute energy demand data from EirGrid for 2014
- Create a datetime column
- Use linear interpolation to fix daylight saving issues
- Plot a time series for energy demand and label the figure accordingly



## Question 2 : Autocorrelation analysis

Autocorrelation refers to the degree of correlation of the same variables between two successive time intervals. It measures how the lagged version of the value of a variable is related to the original version of it in a time series.

### Steps

- Estimate the autocorrelation coefficients for 10 days
- Lags for 10 days:  $10 \times 24 \times 4$
- X\_axis label must be in days



## Question 3 : Normalising timescale

- Create a time of year variable that ranges between 0 and 1
- Show how the demand varies over the course of the year using a graphic.



## Question 4 : Monthly Averaging time series data

- Calculate the average demand for each month of the year
- Provide a bar graph of the monthly average demand



## Question 5 : Daily demand profile

- Calculate the average demand for each hour of the day for the whole year
- Provide a bar graph of the hourly average demand or daily demand profile



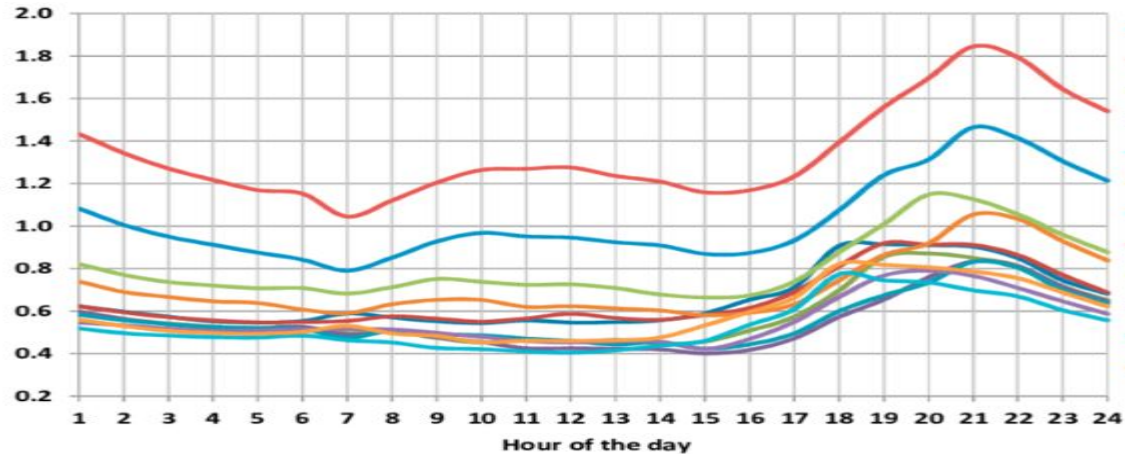


## Question 6 : Weekly demand profile

- Calculate the average demand for each day of the week for the whole year
- Provide a bar graph of the daily average demand

## Question 7 : Variations of the daily demand profiles

- For each hour of the day calculate the average daily demand
- Get daily demand profiles for the 7 days
- Plot the variations of the daily demand profiles over 24 hours





## Question 8: Energy demand between weekend and weekday

- Separate the energy demand for weekends from weekdays
- Perform a statistical hypothesis test to determine whether there is a statistically significant difference between energy demand during the weekend and during the weekday
- Use a significance level  $\alpha = 0.05$

## Question 9 : Benchmark forecasting

- Divide the data into two halves and use the second half for evaluation purposes
- Use the benchmark forecasting approach also known as persistence to forecast the energy demand for the second half of the data
- $\mathbf{\hat{y}(t+k) = y(t)}$
- Horizons for 1 day =  $24 \times 4$
- Calculate the MAE between the predicted energy demand and the actual energy demand of the second half of the data for the forecast horizons with lead times up to one day ahead
- Plot MAE against time leads and interpret the plot



## Question 10 : Benchmark forecasting

Repeat the same process as number 9 instead of using MAE use MAPE

MAE (Mean Absolute Error) is the mean absolute difference between the actual and the predicted value, whilst MAPE (Mean Absolute Percentage Error) is the mean absolute percentage difference between the actual and the predicted value.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_{actual} - y_{predicted}|$$

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_{actual} - y_{predicted}}{y_{actual}} \right|$$

# Submission Files (MATLAB/Python)



- **Single** Python/MATLAB code file(.ipynb or .m) **[Do not Submit checkpoints for .ipynb]**
- Assignment report(.pdf) - remember to name the file as instructed
  - Indicate the libraries you have used in your code at the beginning of the report (After the title page)
- Data files (as given)

## Submission process:

1. Put code **file and data files** in a single folder
2. Name of the folder should be the same as your andrew ID
3. **Zip this folder and attach the zipped file on assignment submission page (CANVAS)**
4. After attaching zipped file, click on "Add Another File" from assignment submission page and **attach your report**
5. Submit your assignment



# Q&A