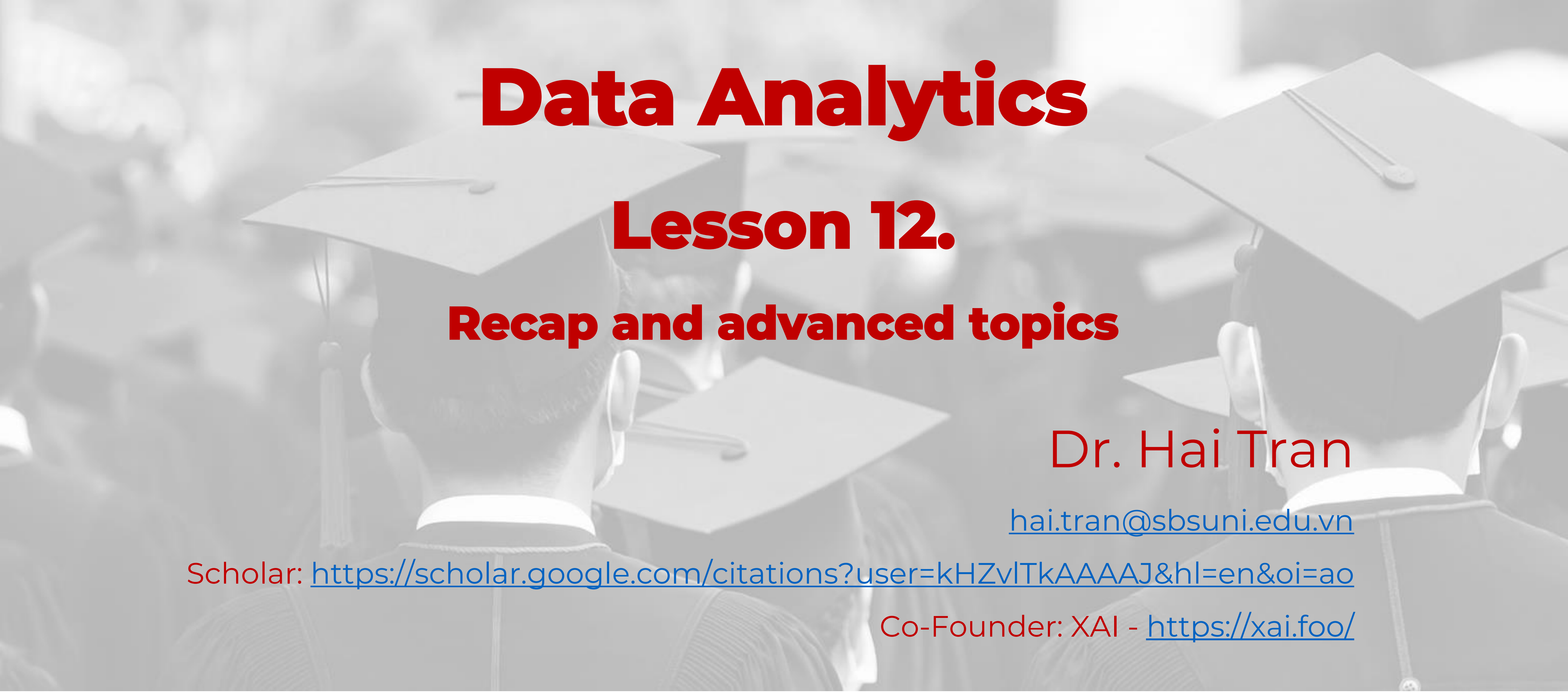


# Data Analytics

# Lesson 12.

## Recap and advanced topics

Dr. Hai Tran

hai.tran@sbsuni.edu.vn

Scholar: https://scholar.google.com/citations?user=kHZvlTkAAAAJ&hl=en&oi=ao

Co-Founder: XAI - https://xai.foo/

SBS Saigon Business School

In partnership with

UCW UNIVERSITY CANADA WEST

MACQUARIE University SYDNEY · AUSTRALIA

UON University of Northampton

AMITY AMITY GLOBAL INSTITUTE

# Learning materials

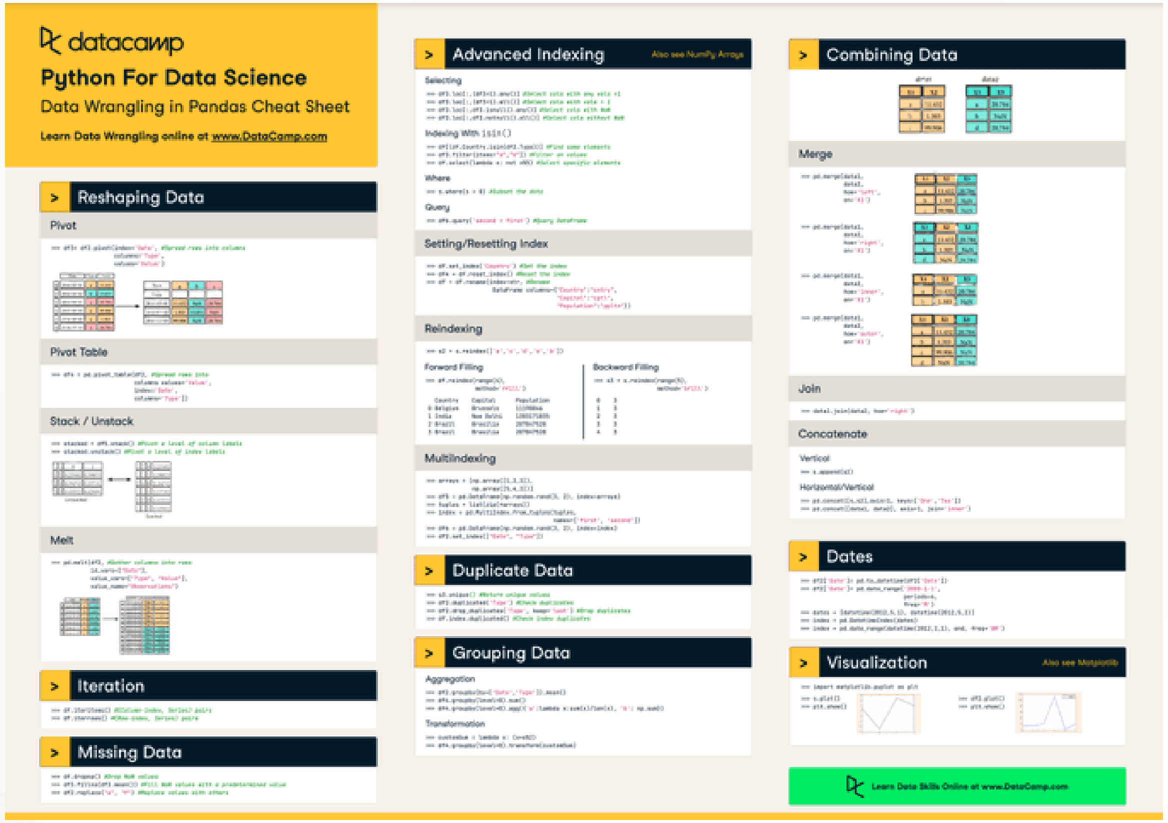
- Textbook
  - Evans, J. (2016) Business Analytics. 2nd edn. Pearson.
  - Runkler, T. (2016) Data Analytics: Models and Algorithms for Intelligent Data Analysis. 2nd edn. Vieweg+Teubner Verlag.
- Online reference materials
  - archive.ics.uci.edu/ml/
  - powerbi.microsoft.com
  - https://github.com/topics/data-analysis-python
  - https://media.pearsoncmg.com/ph/esm/esm_evans_eba3e_20/tools/eba3e_analytic_solver.html
  - https://data.imf.org/

# Agenda

- Lesson 1: Understanding Data Analytics Terminologies.
- Lesson 2: Foundation of Business Analytics
- Lesson 3: Visualizing and Exploring data
- Lesson 4: Applying Descriptive Analytic Techniques
- Lesson 5: Data Modeling
- Lesson 6: Predictive Analytics
- Lesson 7: Regression, Classification and Clustering
- Lesson 8: Forecasting Techniques
- Lesson 9: Investigating Predictive Analytic Techniques
- Lesson 10: Introduction to Data Mining
- Lesson 11: Demonstrating Prescriptive Analytic Methods
- Lesson 12: Recap and advanced topics

Saigon Business School

# Recap and advanced topics



datacamp

## Python For Data Science

Data Wrangling in Pandas Cheat Sheet

Learn Data Wrangling online at www.DataCamp.com

### Reshaping Data

Pivot

Pivot Table

Stack / Unstack

Melt

### Iteration

### Missing Data

### Advanced Indexing

Also see NumPy Arrays

Selecting

Indexing With isin()

Where

Query

Setting/Resetting Index

Reindexing

Forward Filling

Backward Filling

MultiIndexing

### Duplicate Data

### Grouping Data

Aggregation

Transformation

### Combining Data

Merge

Join

Concatenate

Vertical

Horizontal/Vertical

### Dates

### Visualization

Also see Matplotlib

Learn Data Skills Online at www.DataCamp.com

datacamp

## Descriptive Statistics

Cheat Sheet

Learn more online at www.DataCamp.com

### Key Definitions

### Categorical Data—Trail Mix

Counts and Proportions

Visualizing Categorical Variables

Bar plot

Stacked bar chart

Treemap chart

### Numerical Dataset—Glasses of Water

Measures of Center

Other Measures of Location

Measures of Spread

Visualizing Numeric Variables

Histogram

Box plot

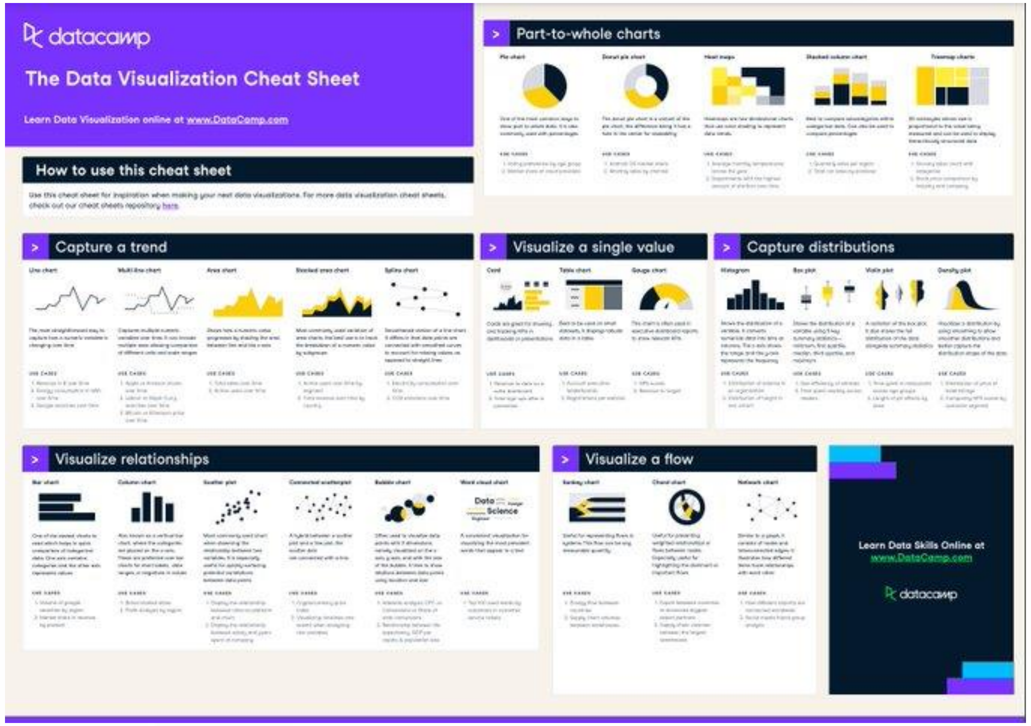
### Correlation

datacamp

Learn Data Skills Online at www.DataCamp.com

Saigon
Business
School

# Recap and advanced topics



datacamp

## The Data Visualization Cheat Sheet

Learn Data Visualization online at www.DataCamp.com

### How to use this cheat sheet

Use this cheat sheet for inspiration when making your next data visualizations. For more data visualization cheat sheets, check out our cheat sheets repository here.

### Part-to-whole charts

Pie chart, Donut pie chart, Heat maps, Stacked column chart, Treemap charts

### Capture a trend

Line chart, Multi-line chart, Area chart, Stacked area chart, Spline chart

### Visualize a single value

Card, Table chart, Gauge chart

### Capture distributions

Histogram, Box plot, Violin plot, Density plot

### Visualize relationships

Bar chart, Column chart, Scatter plot, Connected scatterplot, Bubble chart, Word cloud chart

### Visualize a flow

Sankey chart, Chord chart, Network chart

Learn Data Skills Online at www.DataCamp.com

datacamp

datacamp

## Data Storytelling & Communication Cheat Sheet

Learn more online at www.DataCamp.com

### What is data storytelling?

Data, Visuals, Narrative

### Crafting effective visuals

Choose the best visualization for your story

Line plot, Bar plot, Scatter plot, Histogram

Keep visualizations minimal and avoid clutter

Data visualization decluttering best practices

Use text appropriately

Using text in data visualizations

Use colors effectively

The fundamentals of color theory in data visualization

Do not mislead with data stories

Same Data, Different Y-Axis

Best practices to avoid misleading with data stories

### Crafting effective narratives with data

Know the audience

Executive, Data Leader, Business Partner

Considerations when crafting audience specific messaging

Choose the best medium to share your story

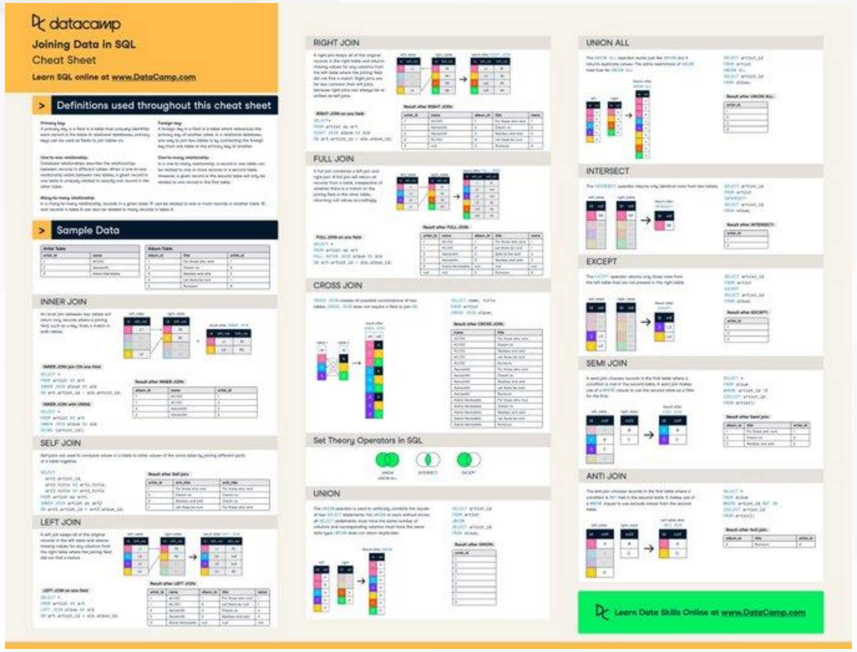
Learn more about data storytelling at www.DataCamp.com

Saigon Business School

# Conclusion & Questions

Here are five advanced topics in data analytics:

1. **Deep Learning and Neural Networks:** Explore advanced techniques in deep learning, including convolutional neural networks (CNNs) and recurrent neural networks (RNNs), to enhance pattern recognition and prediction capabilities.
2. **Natural Language Processing (NLP) and Text Analytics:** Delve into the complexities of analyzing and understanding human language, including sentiment analysis, text summarization, and language generation, to derive insights from unstructured text data.
3. **Time Series Analysis and Forecasting:** Focus on analyzing data that varies over time, such as stock prices, weather patterns, or sales figures, using sophisticated methods like ARIMA (AutoRegressive Integrated Moving Average) and machine learning models for accurate forecasting.
4. **Geospatial Analytics:** Explore the integration of geographic information systems (GIS) with data analytics, allowing for the analysis and visualization of spatial patterns, location-based insights, and geospatial data to make informed decisions.
5. **Explainable AI (XAI):**Address the interpretability and transparency challenges associated with complex machine learning models, ensuring that the decisions made by these models can be understood and trusted by stakeholders, especially in sensitive domains like healthcare and finance.



datacamp

Joining Data in SQL

Cheat Sheet

Learn SQL online at www.DataCamp.com

Definitions used throughout this cheat sheet

Sample Data

INNER JOIN

SELF JOIN

LEFT JOIN

RIGHT JOIN

FULL JOIN

CROSS JOIN

Set Theory Operators in SQL

UNION

UNION ALL

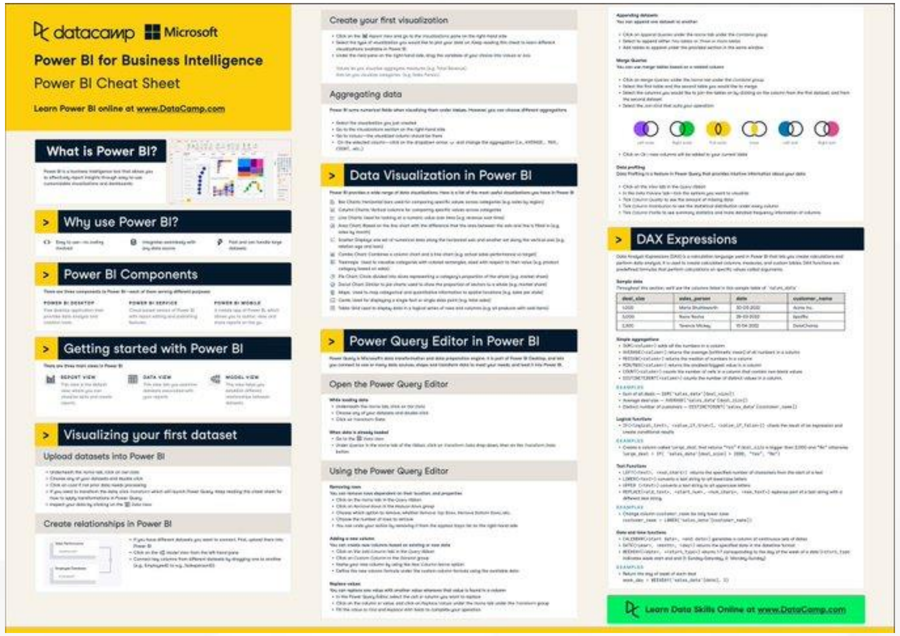
INTERSECT

EXCEPT

SEMI JOIN

ANTI JOIN

Learn Data Skills Online at www.DataCamp.com



datacamp Microsoft

Power BI for Business Intelligence

Power BI Cheat Sheet

Learn Power BI online at www.DataCamp.com

What is Power BI?

Why use Power BI?

Power BI Components

Getting started with Power BI

Visualizing your first dataset

Upload datasets into Power BI

Create relationships in Power BI

Create your first visualization

Aggregating data

Data Visualization in Power BI

Power Query Editor in Power BI

Open the Power Query Editor

Using the Power Query Editor

DAX Expressions

Learn Data Skills Online at www.DataCamp.com

Saigon Business School

# Thank you