

Data Analytics for Management

Week 5: Intro & Excel

Iegor Vyshnevskiy

October 7, 2022



Table of Contents

- 1 Intro
- 2 About the course
- 3 Data Science vs Data Analytics
- 4 Excel

About me

Hello! My name is *Igor*.

- Ph.D candidate KDIS
- MSc KDIS (2014)
- MSc KNEU (2009)
- Background: international finance & central banking
- Research interests: banking and central banking

Contact & profile:

- ievysh@kdis.ac.kr
- Google Scholar
- LinkedIn
- ResearchGate
- GitHub

About you

- What's your preferred name?
- What are you studying / doing?
- What would you like to do?

General things

- Our Goals: to understand concept and tools to further apply in real-life practice
- See that Data Science is *fun* :-)
- My role: to guide you through the course
- Your suggestions are welcomed (through KSS survey, etc.)!

Our communication

- Feel free to approach me before or after a class
- Office hours: by appointment. You are welcome to discuss course-related issues and questions, carrier plans, etc.
- Emailing policy: email me to submit your home work, set up a one-to-one meeting or in the case of some urgent issues.
- Please indicate the course name / section in the subject line and the issue (e.g., [*Data Analytics for Management*] Meeting request).
- Please write at least two times when you would like to meet and a brief description (1-2 sentence) why you want to meet up with me.

Table of Contents

- 1 Intro
- 2 About the course
- 3 Data Science vs Data Analytics
- 4 Excel

Logistics & materials

- Learning by doing
- A typical class will be focusing on practical exercises and projects' discussions
- Please see *the updated syllabus*

Course materials:

- Please see GitHub *Repository*

Broad coverage

- Formal side:
 - Concepts
 - Understanding
 - Real-life cases
 - Problem-solving
 - Application
- Hidden side:
 - Values
 - Communication
 - Cooperation
 - Discussions
 - Standards

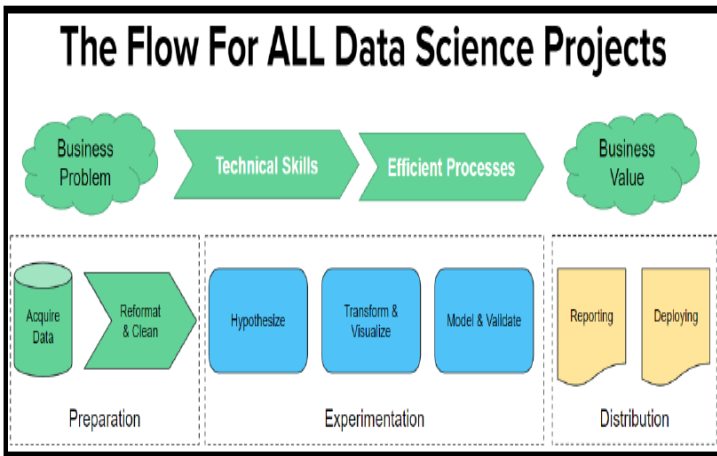
Table of Contents

- 1 Intro
- 2 About the course
- 3 Data Science vs Data Analytics**
- 4 Excel

Concepts

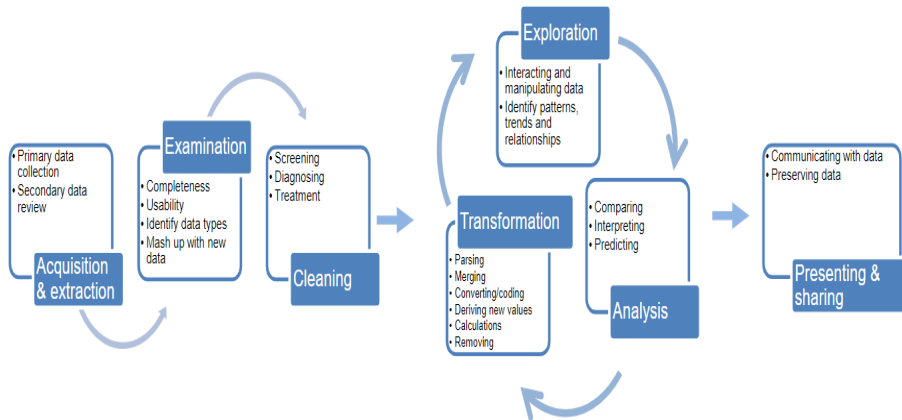
- *Data science* is an interdisciplinary field that uses scientific methods, processes, algorithms and systems to extract or extrapolate knowledge and insights from noisy, structured and unstructured data, and apply knowledge from data across a broad range of application domains. Data science is related to data mining, machine learning and big data (Wiki).
- *Data analytics* (a subfield of data science) is the science of analyzing raw data to make conclusions about that information (Invest).

Data Science workflow¹



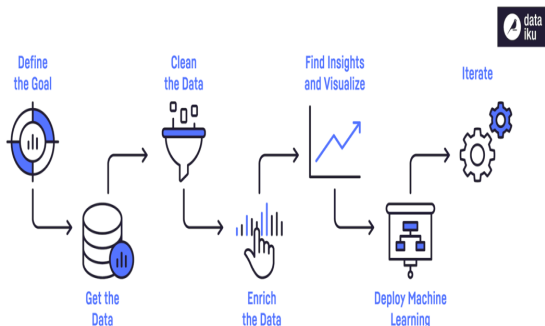
¹Matt Dancho. Everything You Should Already Know About Data Science

Data Analytics workflow (1)²



²Kirk 2012, ACAPS 2013

Data Analytics workflow (2)³



- Specification of Data Requirements
- Data Gathering
- Data Processing
- Data Cleaning
- Data Analysis
- Data Communication

³ A Comprehensive Guide on Microsoft Excel for Data Analysis

Analytics Workflow Stages⁴

- ① Generate: All the ways data is generated and the systems of record where it is stored or originates from, also referred to as data ingress
- ② Collect: All the ways data is collected or ingested
- ③ Prepare: All the ways data is transformed, including ETL (extract, transform, load), ELT (extract, load, transform), reverse ETL (from a warehouse into business applications), and ML (machine learning)
- ④ Store: All the ways data is stored, organized, and secured for analytics purposes
- ⑤ Analyze: All the ways data is analyzed
- ⑥ Deliver: All the ways data is delivered and how it is consumed, also referred to as data egress or data products

⁴Gary Stafford blog. Capturing Data Analytics Workflows and System Requirements

Rules for Data Science⁵

From the board room to the shipping dock, decisions are made every moment of the day using quantifiable, fact-based, trustworthy data (Heine Krog Iversen, Forbes)

- ① Usefulness > Complexity
- ② Data Quality > Hyperparameter Tuning (i.e., set of optimal values)
- ③ Simplicity > Novelty
- ④ Communication > Everything

⁵4 Powerful Rules for Better Data Science

Rules for Data Analysis

One popular view:

Three rules for data analysis:

PLOT THE DATA, PLOT THE DATA, PLOT THE DATA⁶

⁶Uwe H. Kaufmann

Table of Contents

- 1 Intro
- 2 About the course
- 3 Data Science vs Data Analytics
- 4 Excel

Rules for working with Excel⁷

① General:

- Save original data
- Work with copied data
- Save a new version in a new file before updating
- Have a do-tracking (track changes and to-do things)
- Add new columns at the far right of your data
- Check for duplicates

- ② Technically specific^a:
 - One Row of Headings
 - No Empty Columns
 - No Empty Rows
 - All Dates must be in a Single Column
 - Every Unique Data must have its own Column
 - No Totals or Subtotals anywhere in your Table
 - No obstructions around your data

^aDavid Brown

⁷based on practical experience

Data analysis in Excel

① Supplementary materials:

- Video tutorial ([link](#))
- Excel file: Gapminder ([link](#))

② What we do:

- Explore the file. What do we have there (use filter, pivot table)?
- What can be done?
- Illustrations: numerical (i.e., mean, SD) and figures
- Data Communication
- Conclusions

Data dashboards in Excel

Excel dashboards make it easy to perform quick overviews of data reports rather than going through large volumes of data⁸.

- ① Supplementary materials:
 - Video tutorial ([link](#))
 - Excel file: Gapminder ([link](#))
- ② What we do:
 - Use a previous Gapminder file
 - Build a dashboard (bunch of graphs)
 - Data Communication
 - Conclusions

⁸What are Excel Dashboards?

Computer lab learning by doing activity



Project 1 group discussions

