

# Data Session – Spring 2021

Nikolas Breitkopf, Data Advisor  
[nikolas.breitkopf@aalto.fi](mailto:nikolas.breitkopf@aalto.fi)



Aalto University  
School of Business

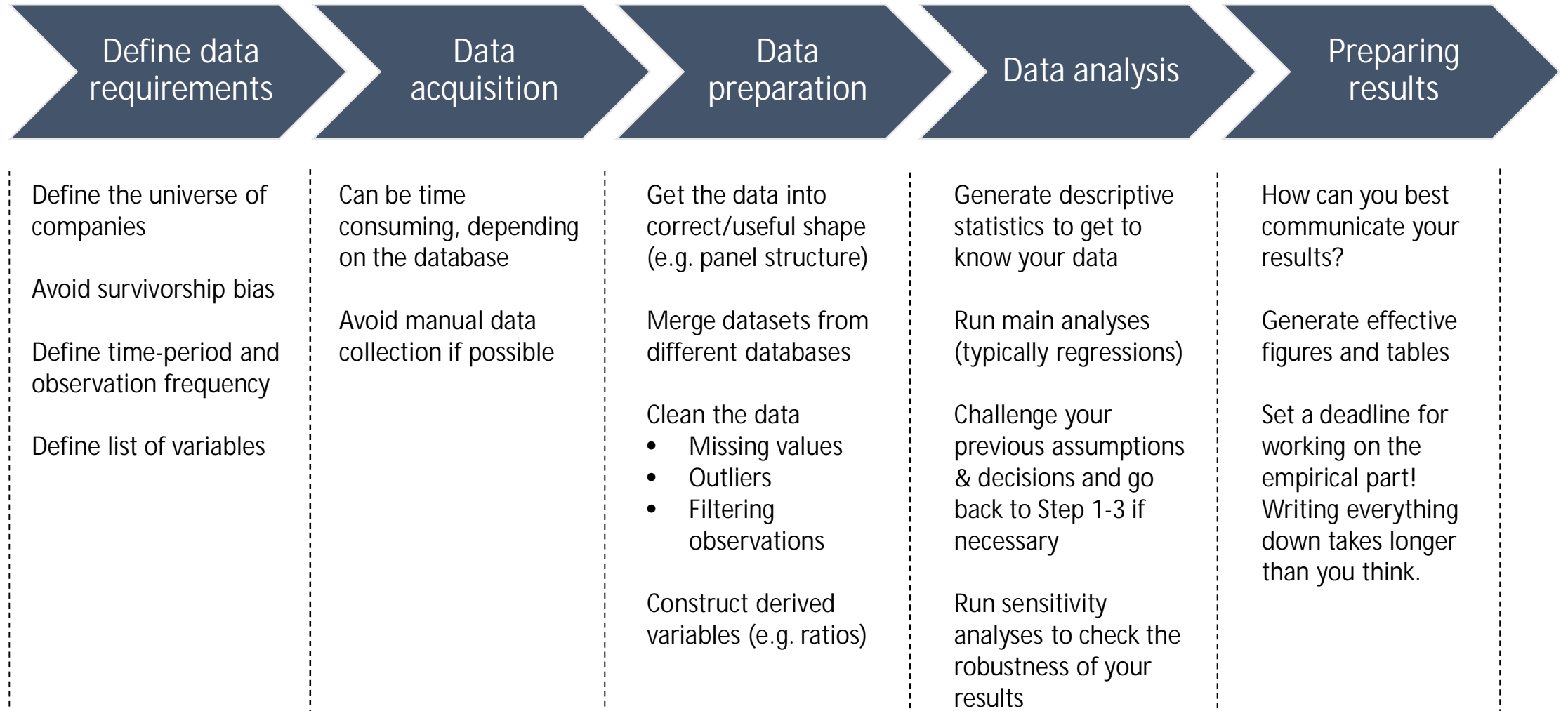
# Agenda

## Part 1:

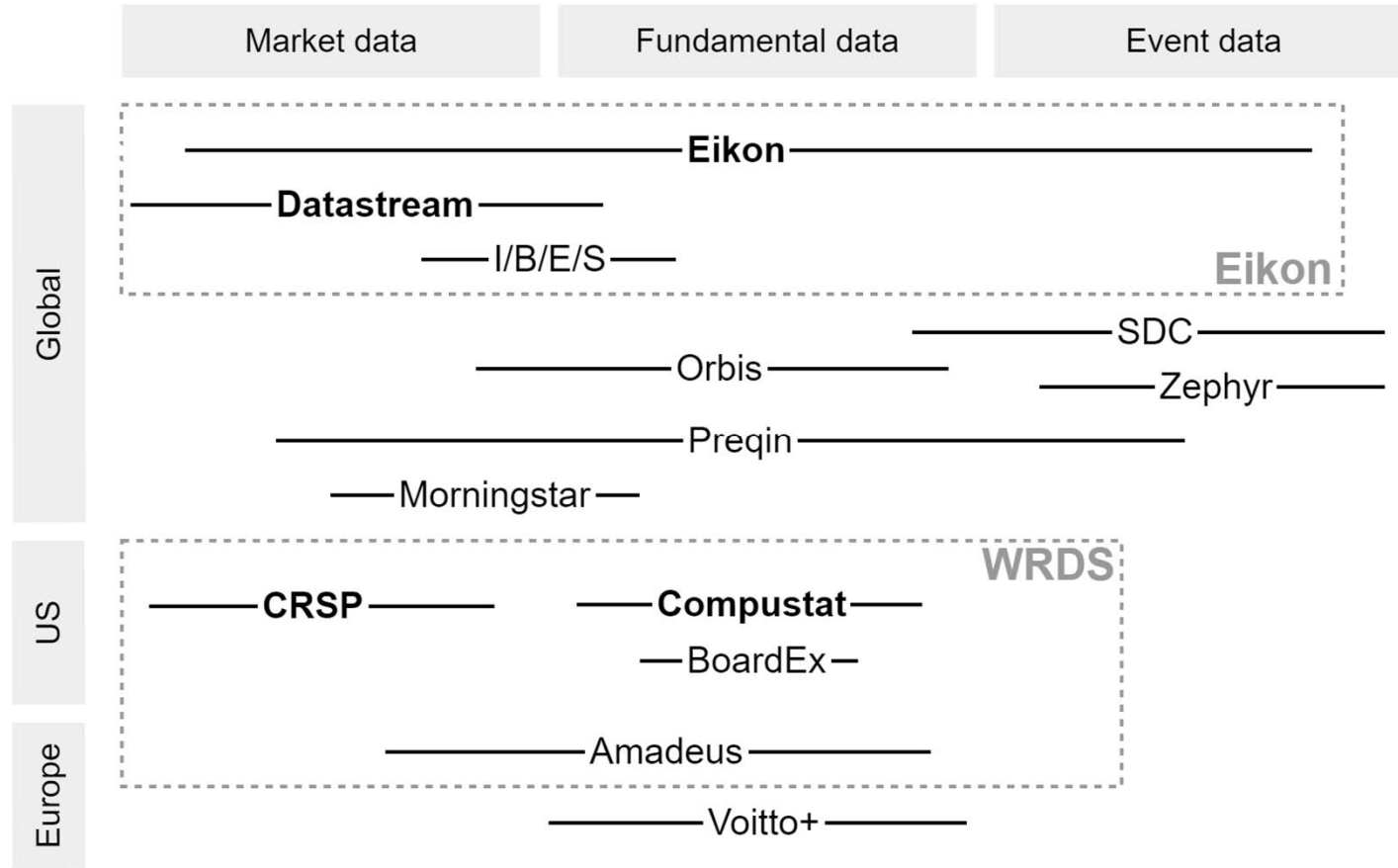
- Workflow of conducting an empirical project
- Which databases are available?
- How to access the databases?
- Demo of common use cases

Part 2: Demonstration of the Preqin database by Anthony Williams (Preqin)

# Workflow of conducting an empirical project



# Overview of available commercial databases



# Main Platforms

- WRDS (Wharton Research Data Services)
  - Web-based platform for downloading research oriented datasets through a common web interface.
  - Convenient access
  - Hosts the most commonly used datasets in Finance research: CRSP/Compustat
- Eikon
  - Information platform by Refinitiv (formerly Thomson Reuters)
  - Very broad in scope
  - Mainly designed for use by the financial industry, but also useful for research
  - Access through desktop application and Excel Add-ins
  - Batch download possibilities sometimes limited

# What is in WRDS?

Database	Data category	Region	Type of firms
CRSP	Market data (security returns, market index returns)	U.S.	public
Compustat	Fundamental data (balance sheet, income statement, cashflow statement, etc.)	U.S.	public
Fama French Factor Returns	Factor portfolio returns, risk-free rates	U.S. Europe	
BoardEx North America	Compensation, employment, and relationship data	U.S.	public and large private
Thomson Reuters Institutional Holdings	Fund holdings, Fund characteristics, ownership structure of public firms	U.S.	mutual funds, investment firms
FINRA TRACE	Bond market data	U.S.	public and large private
Amadeus	Fundamental data (financials, ownership structure)	Europe	private
Preqin	Private Equity and Venture Capital	Global	private

# What is in Eikon?

Data category	Product	EIKON Desktop	Excel Add-on	Python API
Fundamental data (balance sheet, income statement, cashflow statement, etc.)	Refinitiv Fundamentals (formerly Worldscope)	✓	✓	✓
Market data		✓	✓	✓
Corporate actions (Mergers & Acquisitions, IPOs)	Deal Analytics (formerly SDC)	✓	✓ (Deal Analytics Add-on)	✗
Historical time series (equities, equity indices, bonds, credit default swaps, interest rates, macroeconomic time series)	Datastream	✗	✓ (Datastream Add-on)	✗
Analyst Estimates	I/B/E/S	✓	✓	?
Ownership structure		✓	✓	?
Debt structure / ratings		✓	✓	?
ESG scores		✓	✓	?
News		✓	✗	✓

# What is in other commercial databases

Database	Data category	Region	Type of firms
Orbis	Fundamental data (balance sheet, income statement, cashflow statement, ownership structure)	Global	private
Zephyr	Mergers & Acquisitions	Global	public / private
Preqin	Private Equity and Venture Capital	Global	private
Morningstar	Mutual fund performance, characteristics	Global	mutual funds



# Getting access

- Data room, 3rd floor of Väre/Arts building in room Q301
  - Access to Eikon, Datastream, Morningstar, SDC
  - Closed during the pandemic!
- Alternative: Virtual Desktop Infrastructure at <https://mfavdi.aalto.fi/>
  - Remote access to the data room PCs
  - Instructions at <https://wiki.aalto.fi/x/-8kJCw>
- WRDS: Remote access with personal account
  - Instructions at <https://wiki.aalto.fi/x/gsYJCw>
- Heavy Eikon users: Personal installation on own PC
  - Contact me at [nikolas.breitkopf@aalto.fi](mailto:nikolas.breitkopf@aalto.fi)

# Getting access, by database

Database	Virtual Data Room	Browser based	Alternative access options
WRDS		With personal account (see <a href="https://wiki.aalto.fi/x/gSYJCw">https://wiki.aalto.fi/x/gSYJCw</a> )	
Eikon	✓		Personal installation on own PC (for heavy users) Contact me at <a href="mailto:nikolas.breitkopf@aalto.fi">nikolas.breitkopf@aalto.fi</a>
Preqin		With personal account (see <a href="https://wiki.aalto.fi/x/OswJCw">https://wiki.aalto.fi/x/OswJCw</a> )	WRDS
Orbis		Aalto library web-proxy (see <a href="https://wiki.aalto.fi/x/2cwJCw">https://wiki.aalto.fi/x/2cwJCw</a> )	
Zephyr		Aalto library web-proxy (see <a href="https://wiki.aalto.fi/x/2cwJCw">https://wiki.aalto.fi/x/2cwJCw</a> )	
SDC	✓		
Datastream	✓ (Excel Add-on)		Personal installation on own PC (for heavy users) Contact me at <a href="mailto:nikolas.breitkopf@aalto.fi">nikolas.breitkopf@aalto.fi</a>
Morningstar	✓		

# Getting help

- Aalto Finance Database Wiki: <https://wiki.aalto.fi/display/FINDB>
  - General information, how-to articles, FAQs, links to additional resources
- Aalto Finance Database Blog: <https://blogs.aalto.fi/databases>
  - By Henri Blomster (previous Data Officer)
- Product specific documentation on WRDS
  - Manuals with variable descriptions, overviews, FAQs
- Training videos for Eikon: <https://training.refinitiv.com/eikon>
  - Pre-built Eikon templates as starting point
- Office Hours:
  - No fixed office hours during the pandemic. Flexible on-demand virtual communication, instead.
  - If you get stuck, email me at [nikolas.breitkopf@aalto.fi](mailto:nikolas.breitkopf@aalto.fi) with specific questions or to set up a Zoom call.

# Typical sample construction scenarios

Target Sample	Possible selection criterion	Recommended data sources
Public U.S. firms	Firms listed at NYSE, AMEX or Nasdaq	CRSP & Compustat
Public Non-U.S. firms	<ul style="list-style-type: none"><li>• Historical members of a broad stock market index</li><li>• All firms listed at the main stock market of a country</li></ul>	Datastream for market data Eikon for fundamental data
Private firms	All private firms in a country with average historical total assets exceeding some threshold	Orbis



Add additional information from other databases, depending on research question

# Merging databases

General strategies:

- Find a linking table constructed by the database vendor or by other researchers
- Find a reliable common identifier available in both databases
- Avoid linking by name only

Common use cases:

- Linking CRSP and Compustat:
  - CRSP primary security identifier: PERMNO
  - Compustat primary company identifier: GVKEY
  - Historical linking table offered by CRSP/Compustat Merged
- Linking international public firms across databases:
  - ISIN is available as common identifier in most databases and is generally suitable for linking
- Linking SDC (M&A) firms to fundamental data:
  - U.S. sample: <https://github.com/michaelewens/SDC-to-Compustat-Mapping>
  - International sample: Use Ticker in conjunction with country information and name check

# Demo: WRDS

- Goal: Download Compustat fundamental data of companies traded at NYSE, AMEX and NASDAQ. Include links to CRSP securities.
- Solution:
  - Use CRSP/Compustat Merged  $\Rightarrow$  Fundamentals Annual file
  - Filter observations by exchange code in Stata

# Demo: Eikon Desktop & Excel Add-in

- Goal:
  - What data is available in Eikon about a particular company?
  - How can we download the same information in Excel?
- Solution:
  - Search bar and „360 Menu“ in Eikon Desktop
  - Formula Builder and Templates in the Eikon Excel Add-in

# Demo: Datastream

- Goal: Get daily returns of the members of the HEX Index
- Solution:
  - Static requests with constituent lists to get index members
  - Time series request with the list of obtained identifiers



# Demo: Preqin

- Today, at 16.00, by Anthony Williams (Preqin)
- Other Zoom stream