



Measuring the Value of Data

Canadian System of National Accounts
2023



Delivering insight through data for a better Canada



Statistics
Canada

Statistique
Canada

Canada

Options for measurement:

- 3 approaches that national accounts could potentially use:
 - **Market-based:** value is determined based on the market price of comparable products on the market
 - **Income-based:** value is determined by estimating the future cash flows that can be derived from the data
 - **Cost-based:** value is determined by how much it costs to produce the data

Options for measurement: market-based approach

- Conceptually preferable method to estimate capital investment, but not always feasible
 - Data may be of most value to the business that collects it and it is never sold
 - Price depends on the use/user, and the use can depend on what is observed
- If sold, the data has generally undergone transformation and is bundled with other services
 - 3rd party data is sold after the user's data has been processed (e.g. organizing, cleaning)
- How would repeated sales of same data be measured?

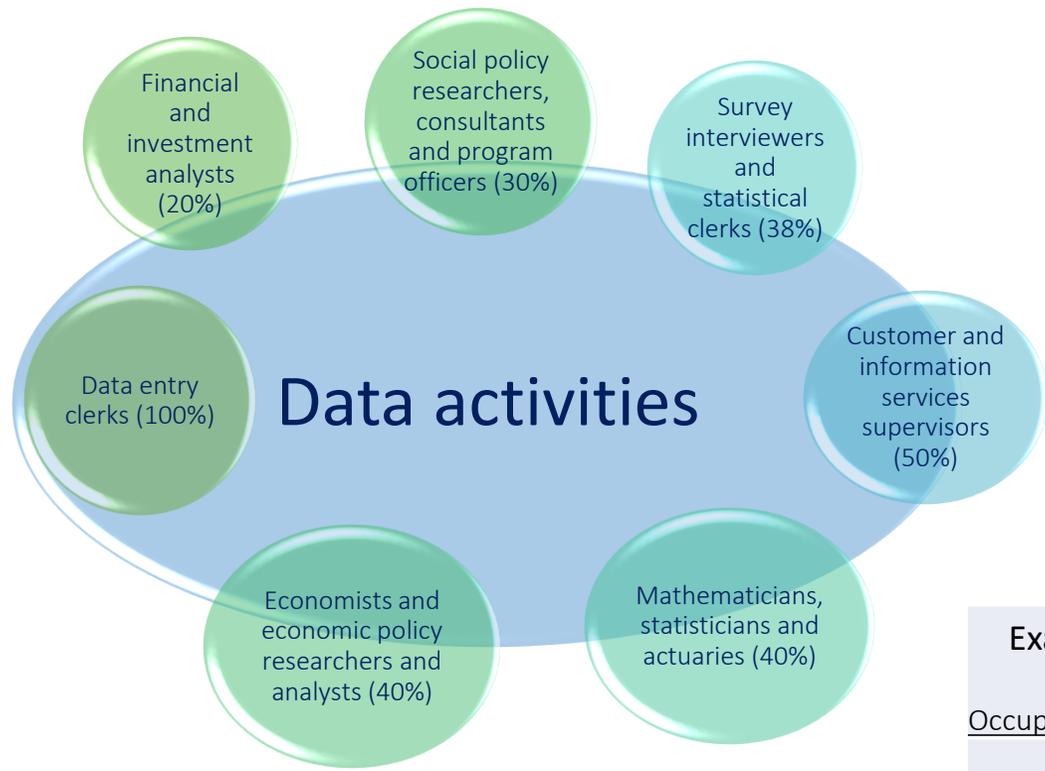
Options for measurement: income-based approach

- Although income-based valuation is an acceptable method, *SNA advises caution*
 - appropriate assumptions about the asset's life length and future cash flows and the discount factor may be difficult to determine
- Often hard to distinguish cash flows (net of associated costs) uniquely related to the data asset from the cash flows related to other intangibles and services
- Income-based approach is recommended for valuing musical, literary, and photographic works— industries where there is an established system of royalty flows

Options for measurement: cost-based approach

- Sum of costs approach is the recommended method in absence of observable market transactions and for own-account production
- Includes an estimate of labour costs, indirect costs and capital services
 - Labour costs = # of employees * average compensation * average time spent
 - Indirect costs include HR resources, electricity, building maintenance, etc.
 - Capital services represents the return on capital assets used in this productive activity

Sum of costs approach to value data activities



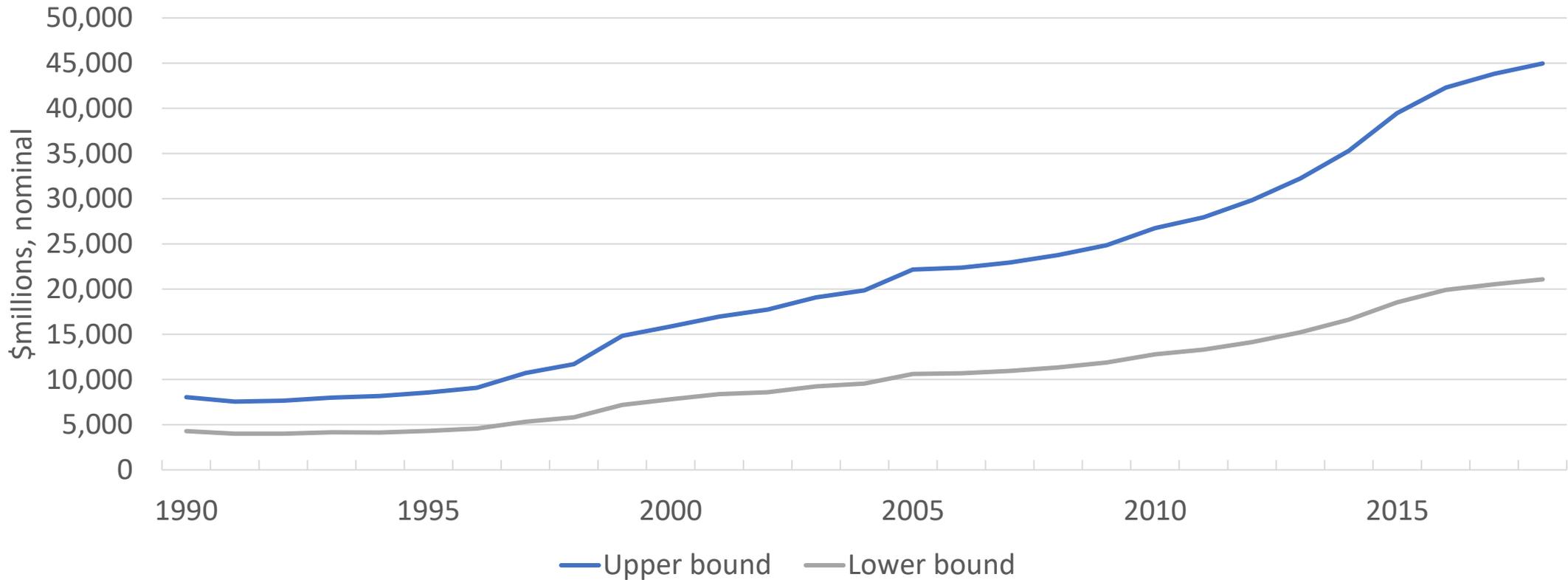
- What occupations are involved in data activities?
 - What portion of their tasks relate to data?
 - What should be the markup to cover non-direct salary costs?
- *Apply to the wage bill to estimate investment*

Example:	'Data' share of production activities	Markup for non-direct-salary costs	Labour compensation	Investment in 'DATA'
<u>Occupational group</u>				
			(\$millions)	
Financial and investment analysts	20%	53%	7,348	2,249
Customer and information services supervisors	50%	53%	668	511



Investment in data activities

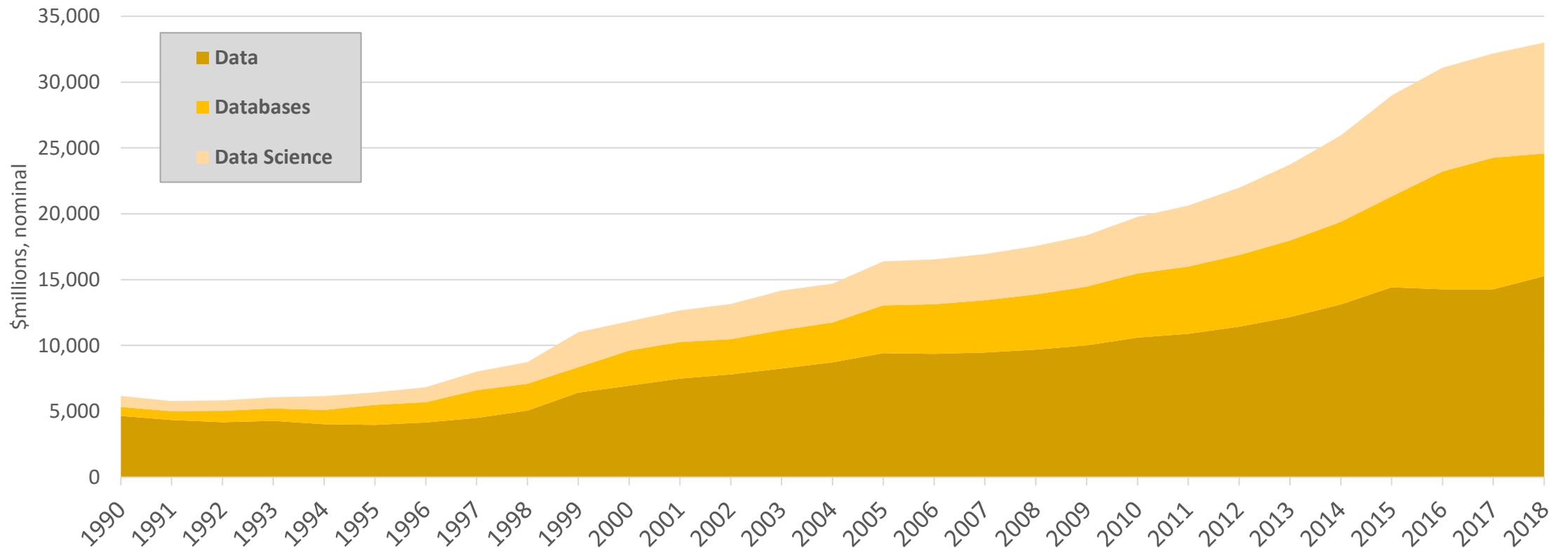
Upper and lower ranges





Investment, by type of Data asset

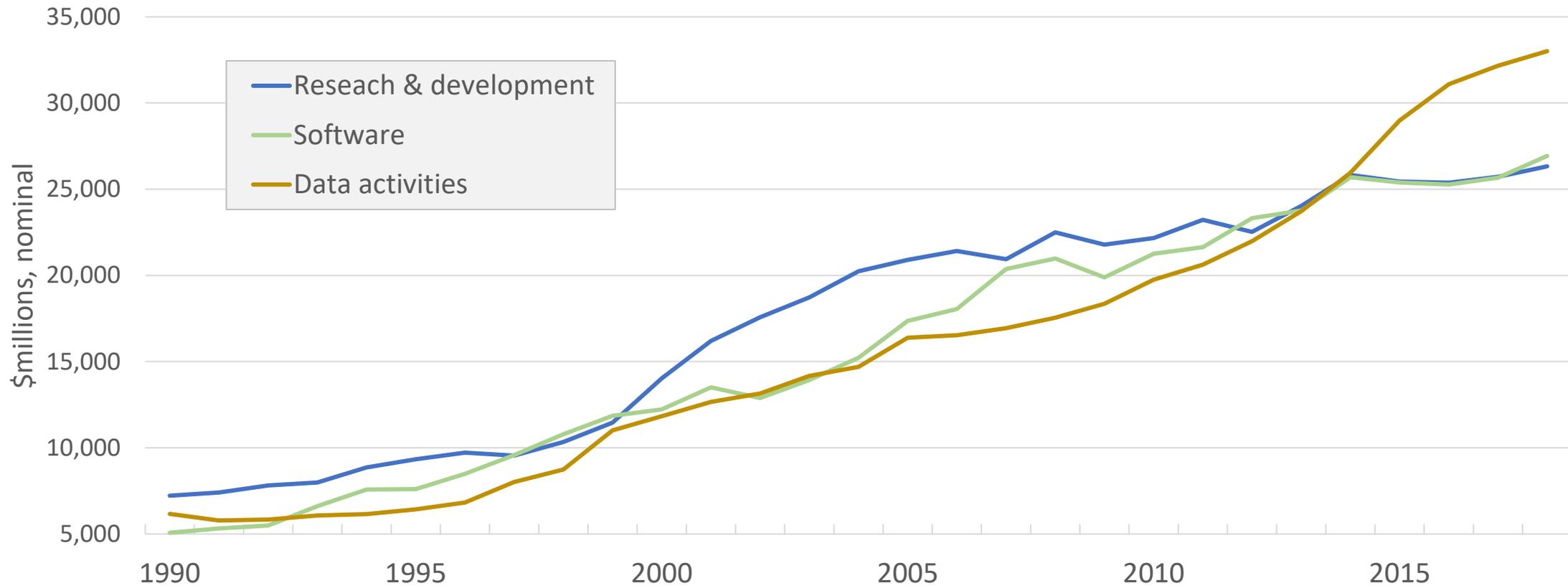
Investment Flows



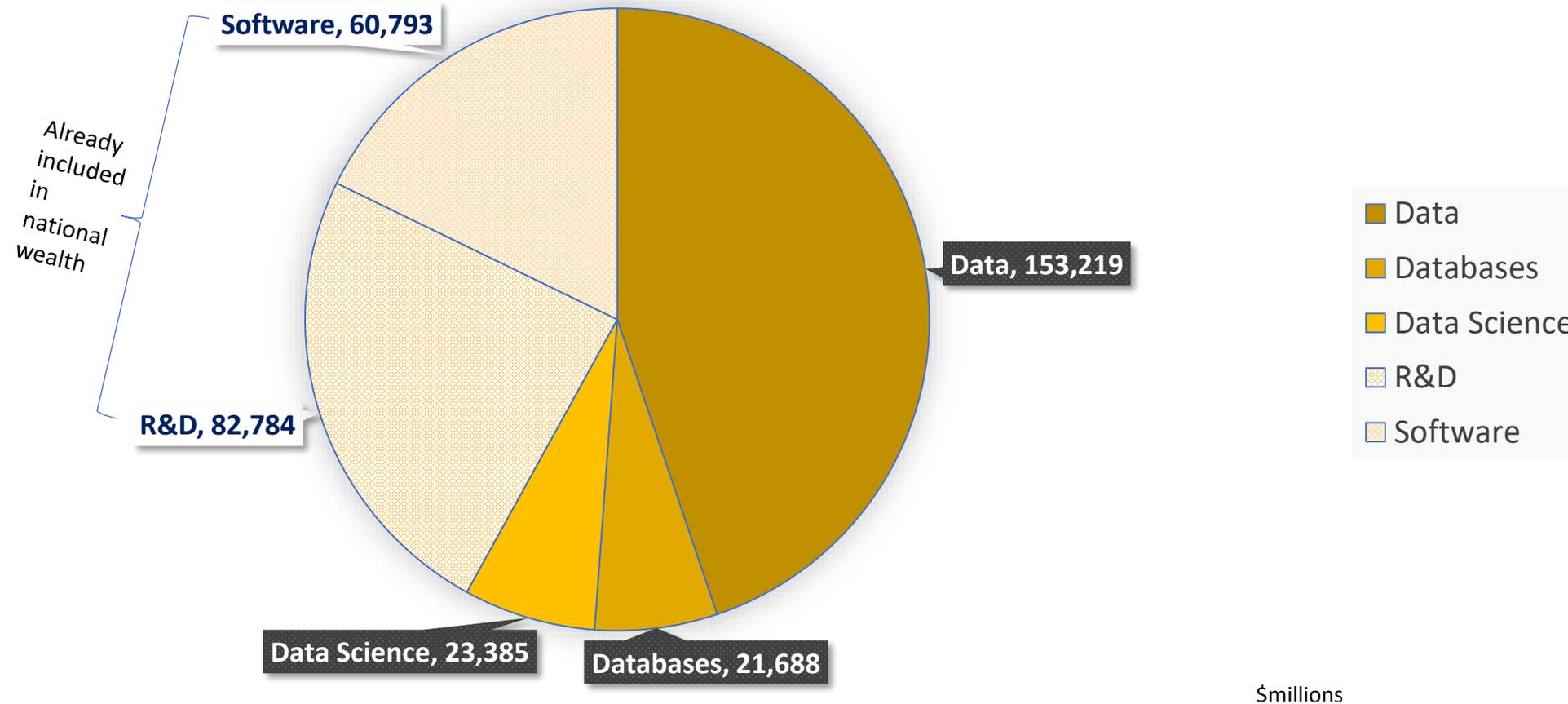


Investment in intangibles: with data activities

Investment in R&D, software and data activities



What is the stock of data activities?



Sum of costs approach: considerations

- Overlap with other IPP production:
 - “it is important to ensure, in using the sum of costs approach to valuating of IPP assets produced on own-account, that the same costs are not included in the valuation of more than one asset”
 - we know there is some overlap with software
- Choice of occupation and time spent on these activities:
 - This could be very narrow or very broad
 - Concrete examples of what is considered Data activities is required
- Mark-up to cover indirect costs and capital services

Update to the Value of Data Activities release

- Estimated based on the Sum of Costs method
 - Certain occupation types, estimates of time spent
- *Previous methodology was ad hoc/arbitrary*
 - *which occupations to include and time-use proportions*
- Challenges related to estimating occupations and time:
 - Occupations engaged in data-related tasks are not obvious
 - Tasks are evolving as our economy evolves
 - Occupations may be involved in more than one stage

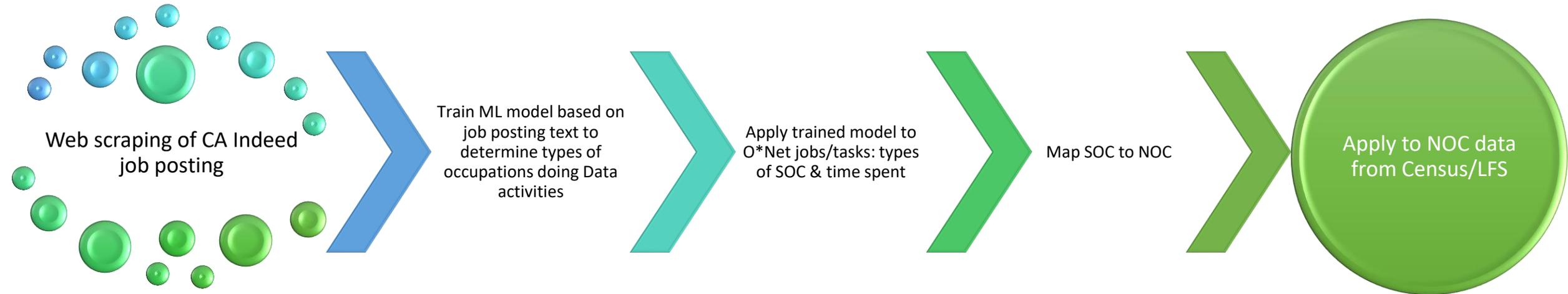


Update to the Value of Data Activities release, cont.

- Use machine learning to refine the types of occupations and time-use
 - Web scraping to obtain jobs listings
 - Machine learning to identify occupations involved in data activities based on key words
 - Linking those occupations to O*Net from the BLS that lists job tasks and importance



Update to the Value of Data Activities release, cont.



*SOC: US occupational classification

Update to the Value of Data Activities release

- This will give a selection of occupations that are participating in data activities based on ‘real’ tasks
- Directly comparable with the US estimates

Can this method be used for other Sum of costs estimation?

- Considerations with this method:
 - only one job posting site
 - set period of time
 - US listing of occupation tasks/importance

