

Survey Design 101

Part II: Analyzing and Interpreting Survey Findings



Okan Bulut, Ph.D.

Centre for Research in Applied Measurement and Evaluation


University of Alberta



Hello!

I am **Okan Bulut**

Associate Professor in Measurement, Evaluation, and Data Science ([MEDS](#)) & Program Coordinator

You can find me (virtually) at:  bulut@ualberta.ca

 [@drokanbulut](https://twitter.com/drokanbulut)

 www.okanbulut.com



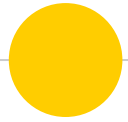
Agenda

1. Analyzing survey data
2. Interpreting survey findings

1

Analyzing survey data

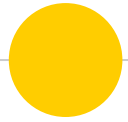
Data analysis options based on the type of survey data



How to analyze survey data



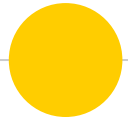
- ① What type of variables (quantitative or qualitative) do you have in your survey data?
- ① What are the possible options for analyzing survey data?
 - Statistical and qualitative data analysis
 - Software programs (open-access vs. commercial)
- ① Who is the target audience for the findings of your survey?
- ① How do you plan to present your findings?
 - Thesis/dissertation
 - Publications, technical reports
 - Posters and presentations



Types of variables

Surveys typically yield two types of data:

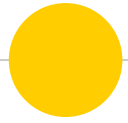
- **Quantitative data (*less common*)**
 - Age, income, height, weight
 - Time-related variables (e.g., years of experience)
 - Counts of things (e.g., children, books, credits earned)
- **Qualitative data (*more common*)**
 - Nominal categories (gender, race, province of residence)
 - Ordered categorical (level of satisfaction or agreement, frequency of a behaviour)
 - Descriptive comments (e.g., open-ended questions)



Descriptive statistics

Some descriptive statistics that we can report for survey data include:

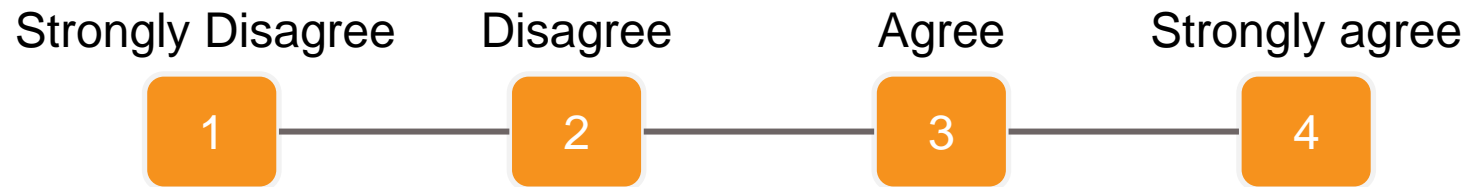
- 1. For quantitative variables:**
 - Mean, median, standard deviation, minimum, and maximum
- 2. For qualitative variables (either ordinal or nominal):**
 - Nominal variables: Frequencies and proportions (i.e., percentages)
 - Ordinal variables: Frequencies, proportions, median, and mode



Avoid: Averaging ordinal variables

Ordinal variables such as Likert-scale questions should **not** be averaged or summed.

Expectation:



Reality:





Item and scale analysis

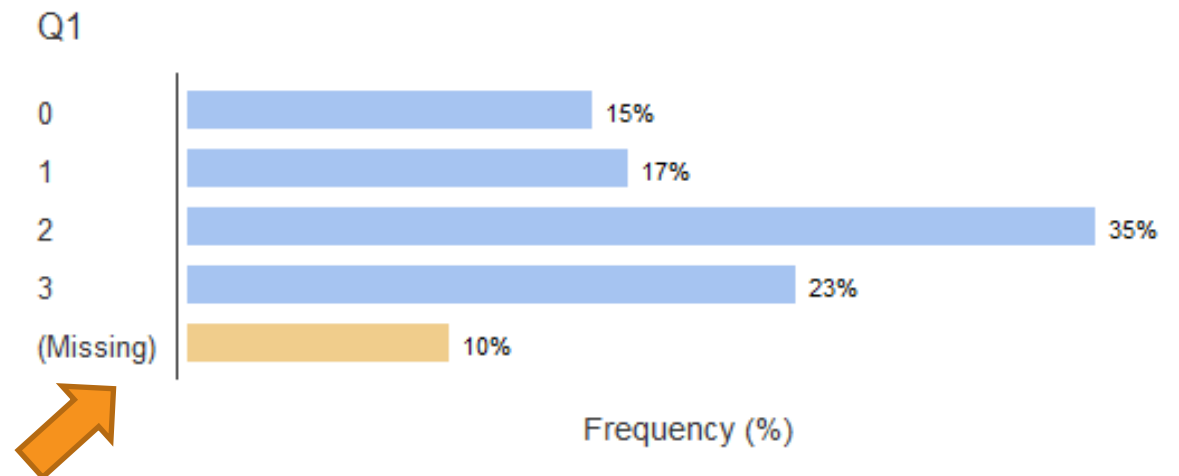
1. Check response frequencies

- Response categories (enough utilization of each response option)
- Skewness in the responses (no extreme skewness is desired)
- Levels of missing data (should be less than 10%)

Q1: I enjoy learning mathematics.

Response options are:

- 0 indicating "I strongly disagree"
- 1 indicating "I disagree"
- 2 indicating "I agree"
- 3 indicating "I strongly agree"





Item and scale analysis

2. Check the alignment among the questions

Questions focusing on “enjoying math” on the survey

Q1: I enjoy learning mathematics.

Q2: I wish have not to study Math. (reverse-coded)

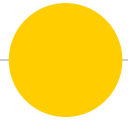
Q3: Mathematics is boring. (reverse-coded)

Q4: I learn interesting things in mathematics class.

Q5: I like mathematics.

Q6: I think it's important to do well in mathematics.





Item and scale analysis

3. Check scale reliability (for a group of questions)

Questions focusing on “enjoying math” on the survey

- Q1: I enjoy learning mathematics.
- Q2: I wish have not to study Math. (reverse-coded)
- Q3: Mathematics is boring. (reverse-coded)
- Q4: I learn interesting things in mathematics class.
- Q5: I like mathematics.
- Q6: I think it’s important to do well in mathematics.

Scale Reliability Statistics

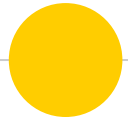
Cronbach's α	
scale	0.869

should be > 0.70

Item Reliability Statistics

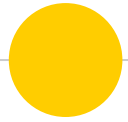
	item-rest correlation	if item dropped
		Cronbach's α
Q1	0.840	0.815
Q2	0.678	0.846
Q3	0.681	0.845
Q4	0.618	0.855
Q5	0.817	0.818
Q6	0.372	0.887

should be > 0.20



Analyzing qualitative responses

- ⦿ Thematic analysis of open-ended responses
 - Generating initial codes or labels
 - Searching for themes or common patterns
 - Defining and naming each theme
- ⦿ Word clouds using open-ended responses
- ⦿ Text mining analysis
 - Sentiment analysis to extract emotions from the text
 - Topic modeling to create themes automatically



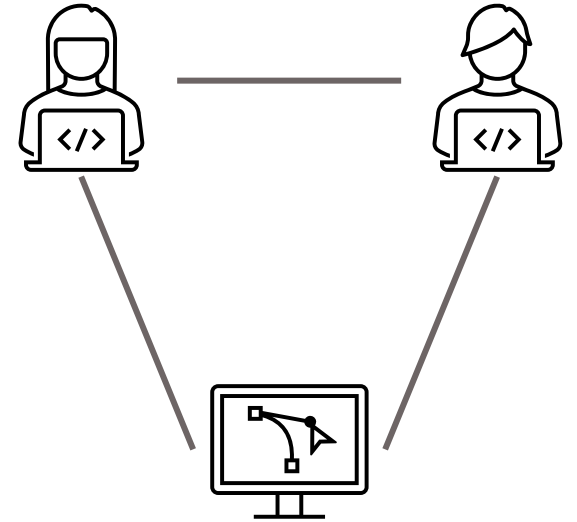
Statistical software programs

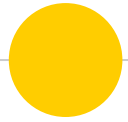
Open-access (free) programs

- jamovi - <https://www.jamovi.org/>
- JASP - <https://jasp-stats.org/>
- PSPP - <https://www.gnu.org/software/pspp/>
- R (no GUI) - <https://cran.r-project.org/>

Commercial programs

- Microsoft Excel
- IBM SPSS Statistics
- Minitab





Software programs for qualitative data

Important: Qualitative software programs facilitate data storage, coding, retrieval, comparing, and linking; but they do **not** analyze data (Patton, 2015, p. 529)

- ◎ Open-access programs
 - [Compendium](#) – for mapping and management of ideas and arguments
 - [KH Coder](#) – for content analysis and text mining
 - R and Python (no GUI)

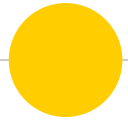
- ◎ Commercial programs
 - [MAXQDA](#), [ATLAS.ti](#), and [NVIVO](#)
 - Check out [UALberta On the Hub](#) for discounted prices
 - [discovertext](#) for analyzing textual data



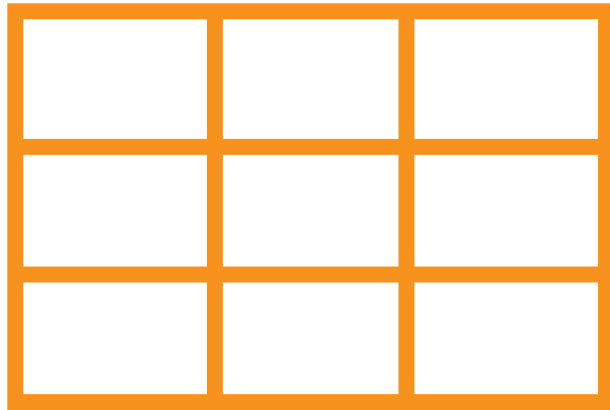
2

Interpreting survey findings

How to interpret and report survey findings



Presenting survey findings



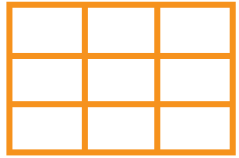
Tables



Figures



Tables vs. Figures



- Prefer **tables** if:

- The table will be used to look up **individual** values
- **Precise** information is required
- Both **summary** and **original** values are needed



- Prefer **figures** if:

- The goal is to reveal relationships among whole sets of values
- The relationships between two or more variables are the main priority



Creating effective tables

Table 1. **Median** scores for items on the **Social Engagement Scale** by students' pass/fail status in mathematics.

As a 9 th grade student, I ...	Fail	Pass	Total
care if I finish high school.	4	4	4
know my way around the school.	4	3	3
have many friends at school.	3	4	3
know what my teachers think about me.	2	3	3
ignore what other students think about me.*	2	2	2
come to school only because my friends are here	2	2	2

Note: Students failing mathematics $n = 31$; students passing mathematics $n = 147$. Response scale for the questions: 1 = Strongly disagree; 2 = Disagree; 3 = Agree; and 4 = Strongly Agree.

* The question uses **negative wording**, and a low median indicates positive student behaviour.



Avoid: Tables with a single row

	Telephone Survey		City Insight Panel	
	Total Satisfied (4,5 ratings)			
	2014	2015	2014	2015
	n=400	n=400	n=1,126	n=1,630
Overall satisfaction with living in the City of Edmonton	77%	80%	75%	74%

Source: 2015 Edmonton Citizen Perception Survey

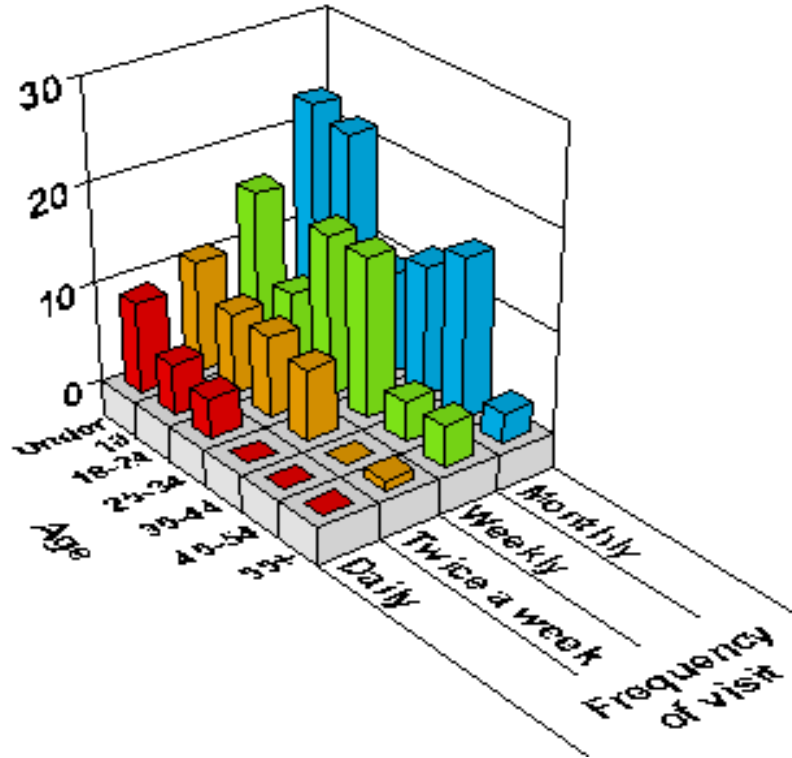
● Creating effective figures

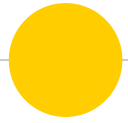
- Figures can be very effective for summarizing survey findings.
- Consider the following elements
 - Colours
 - Shapes
 - Size
 - Font type, size, and colour
- Check out [my workshop notes](#) on visualizing survey items





Avoid: Complex figures





Less is **more...**

Remove
to improve
(the **data-ink** ratio)

Created by **Darkhorse Analytics**

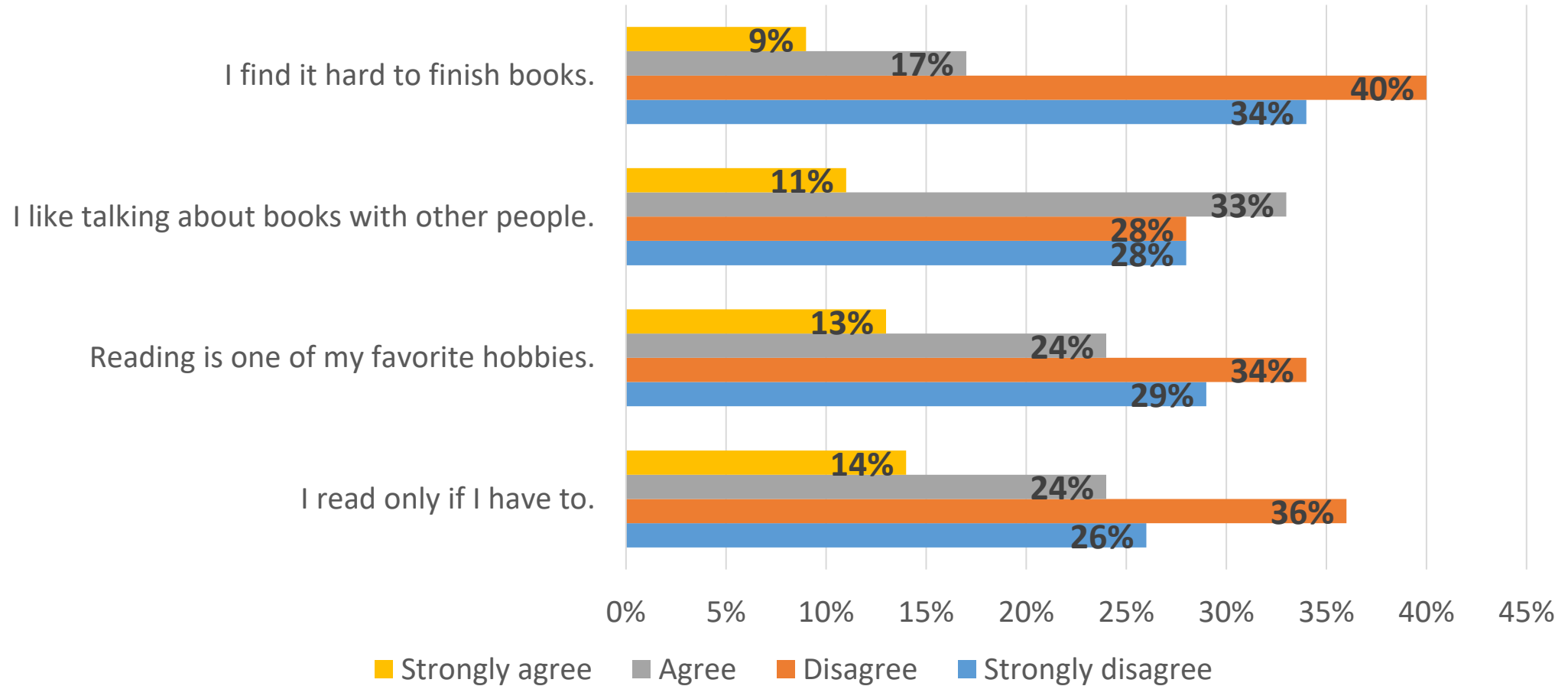
www.darkhorseanalytics.com

Source: [Darkhorse Analytics – Data looks better naked](#)



Bar charts

PISA 2009 Results (Source: OECD)

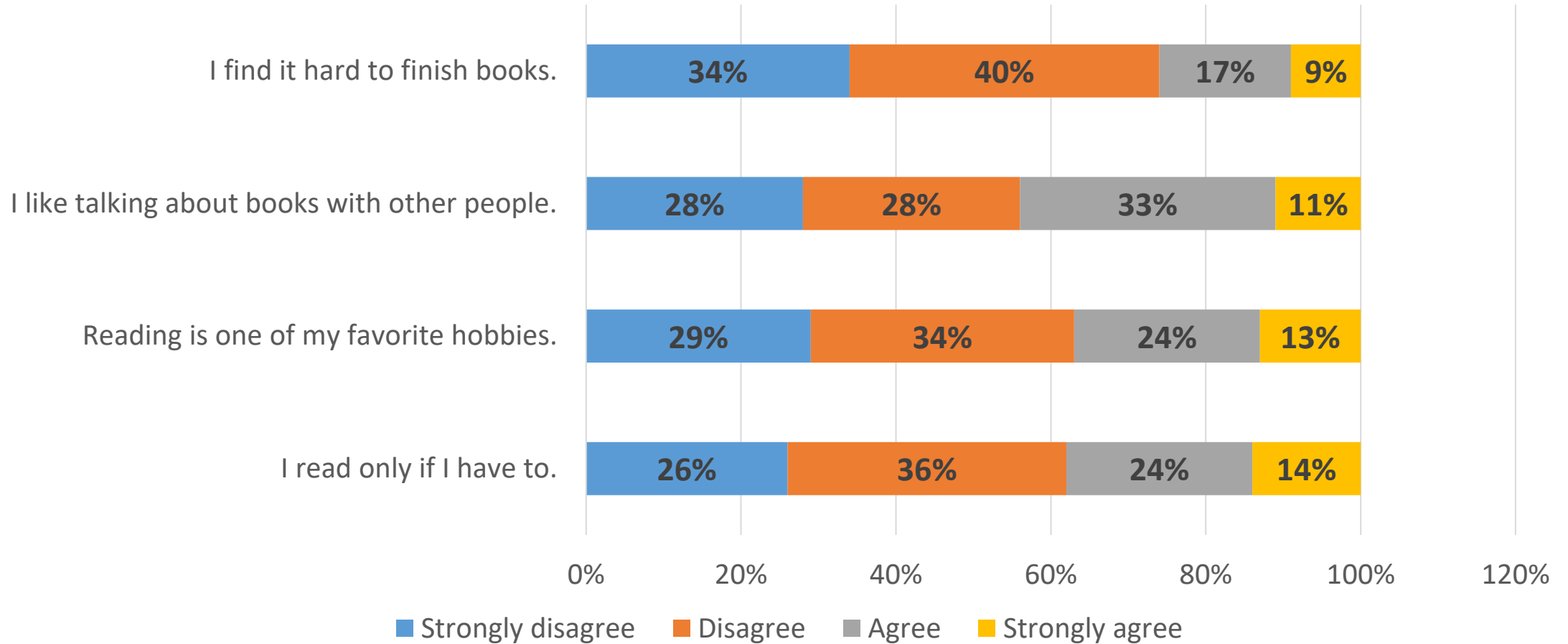


Note: The data come from a random sample of Canadian students who participated in PISA 2009.



Bar charts

PISA 2009 Results (Source: OECD)



Note: The data come from a random sample of Canadian students who participated in PISA 2009.



“Some” tools to make great figures

- Microsoft Excel:
 - Bar charts, pie charts, scatter plots, line charts, etc.
 - Excel gives publication-quality graphs (a bit tedious...)
 - <https://www.youtube.com/watch?v=d7JgmQCLgNg>
- jamovi
 - Bar charts, scatterplots, boxplots, correlation matrix plots, etc.
 - Not very customizable but high-quality graphs
 - <https://www.youtube.com/watch?v=NoG-YJcTtx8>
- Websites to create graphs
 - **RAWGraphs** (<https://rawgraphs.io/>): Free and easy to use
 - **Datawrapper** (<https://www.datawrapper.de/>): Free plan; charts, tables, and maps
 - **Canva Graph Maker** (<https://www.canva.com/graphs/>): Beautiful infographics; values must be entered manually.



Thanks!

Any **questions?**