Betty Harper, Shannon Lee, and Carly Sunseri

The Pennsylvania State University

Build it right the first time: Modeling survey data for BI tools
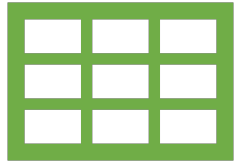
Welcome!
We are…

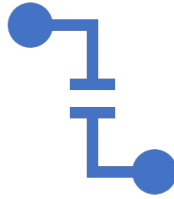Betty Harper, Assistant Director of Assessment

Shannon Lee, Research Associate

Carly Sunseri, Data and Reporting Analyst

The Office of Planning, Assessment, and Institutional Research (OPAIR), The Pennsylvania State University

By the end of this session, you should be able to:

Describe flat files and their limitations

Sketch out a basic relational data model

List the key steps to build a data model

Find additional resources for self-guided learning

...so that you can get started on your survey data visualization journey

Session Roadmap

- Flat files and their limitations
- Facts and dimensions
- Building your base files in Excel
- Building the data model
- Conclusion

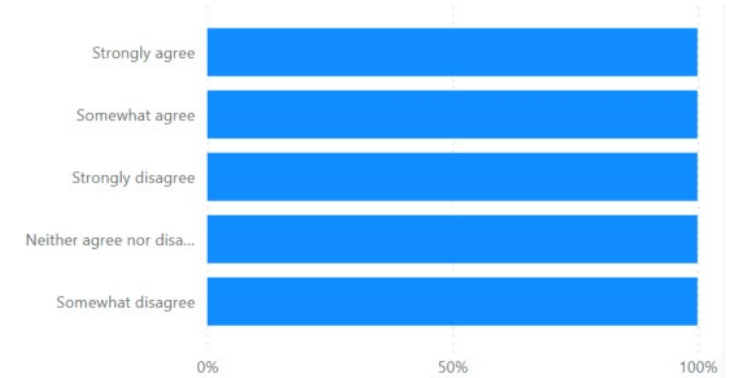# Building Blocks

The initial steps in your journey

Flat files

Why a Data Model?

Facts vs. dimensions

Facts = things that can be measured such as multiple-choice questions

Dimensions = companion data such as demographics

| To what extent do you agree with the following statements?	
Q1	I feel welcome at Faber College.
Q2	I feel valued at Faber College.
Q3	I feel safe at Faber College.
Q4	I think fraternities are an important part of Faber College student life.
Q5	I think togas are appropriate classroom attire.
Q6	I do not appreciate the pop culture references in this survey.
Please answer the following questions.	
Q7	How often do you skip class to attend fraternity parties?
Q8	How often do you take advantage of Dean Wormer's office hours?
Q9	How would you rate the overall quality of your education Faber College?
Q10	Would you recommend Faber College to others?

Q11	Gender identity
Q12	Age range
Q13	IPEDS Race

Where do I start?

Fields
- Unique ID
- Q1 – Q10
- Gender
- Age Range
- Race/Ethnicity
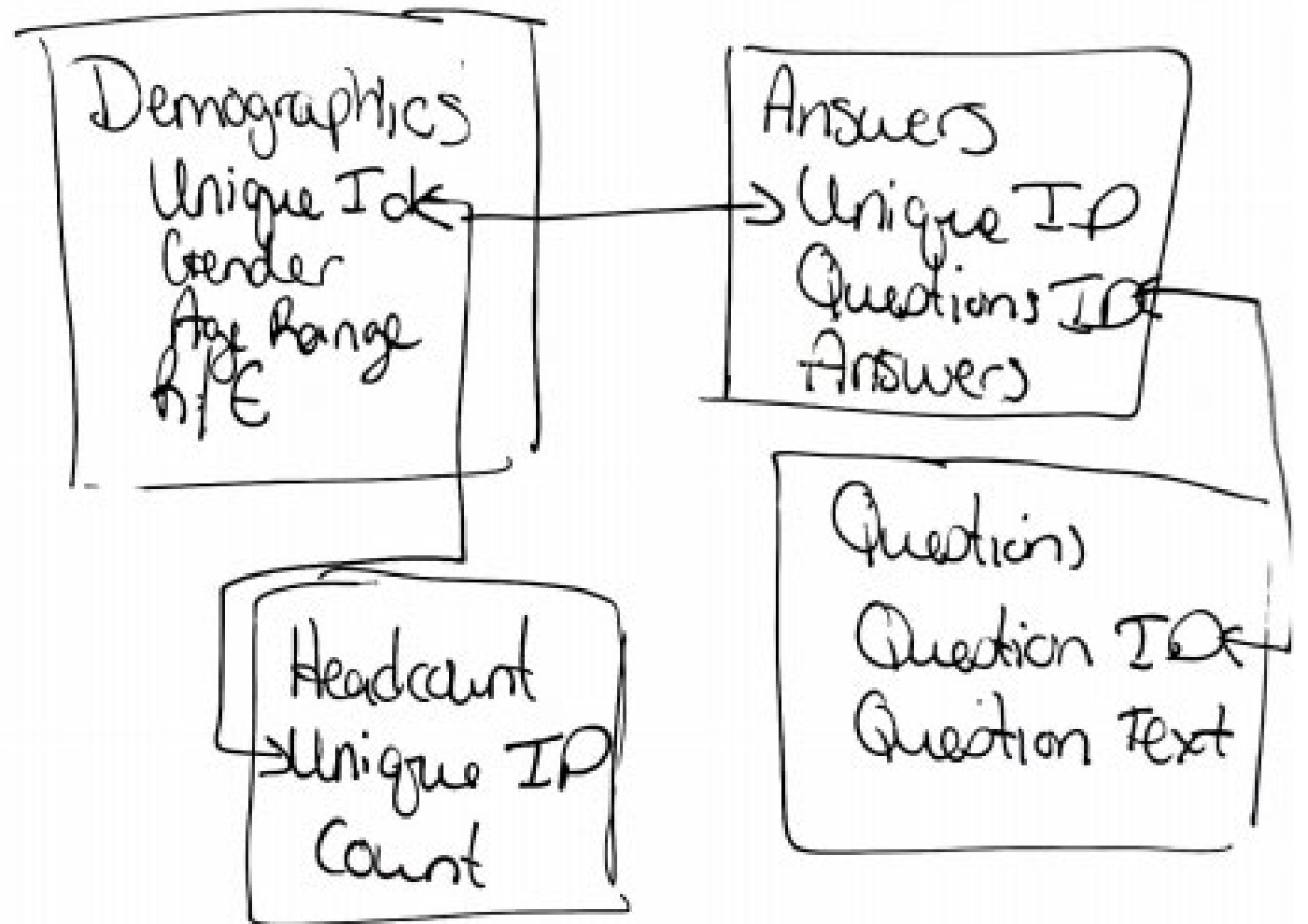
Creating your base files in Excel

- Answers
- Question Key
- Headcount
- Demographics
- Value Order

*Your survey data may result in more files than this
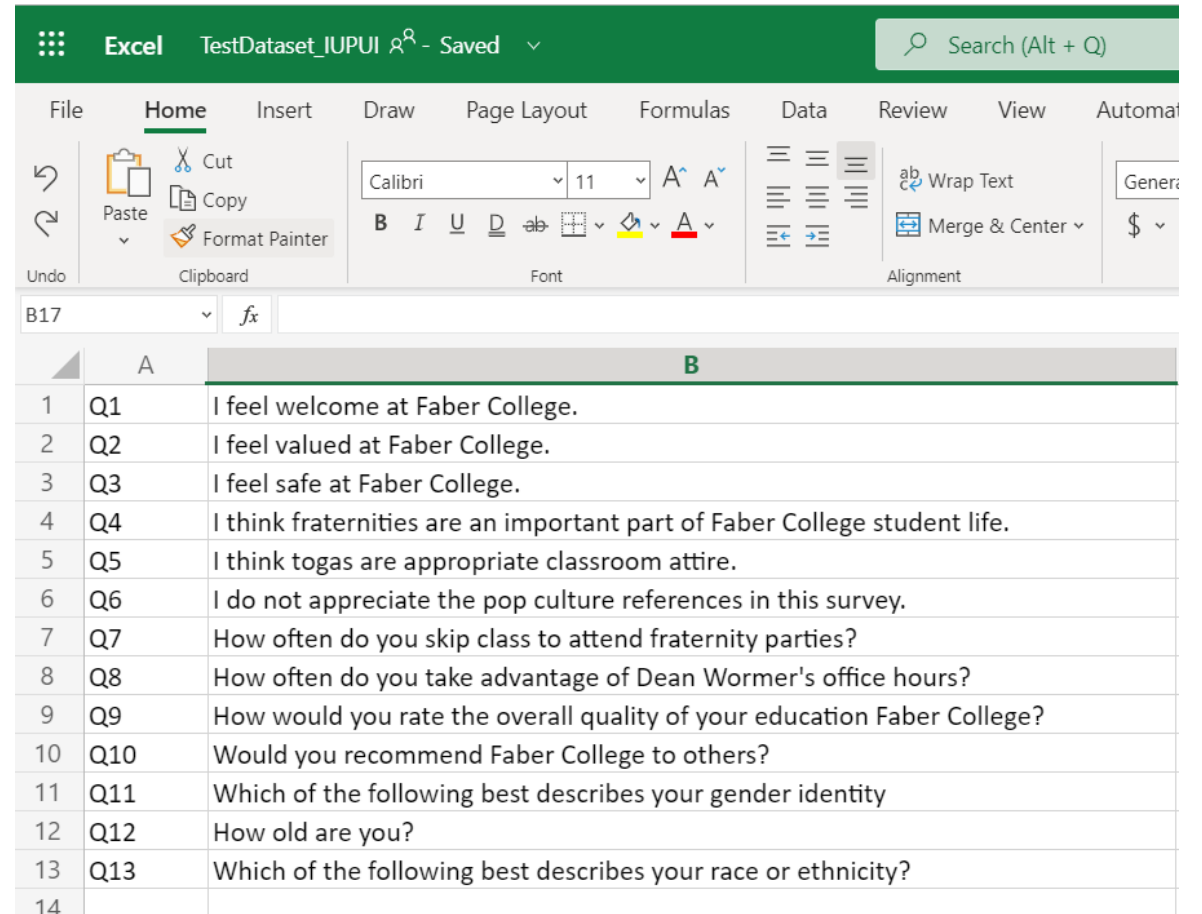
Survey answers (facts)

The question key

Headcount

Demographics (dimensions)

Value Order

(used to make sure responses are presented in your desired order later)

Importing your data into your visualization tool

We will use Power BI to demonstrate

Importing Your Data

Power BI makes it easy!
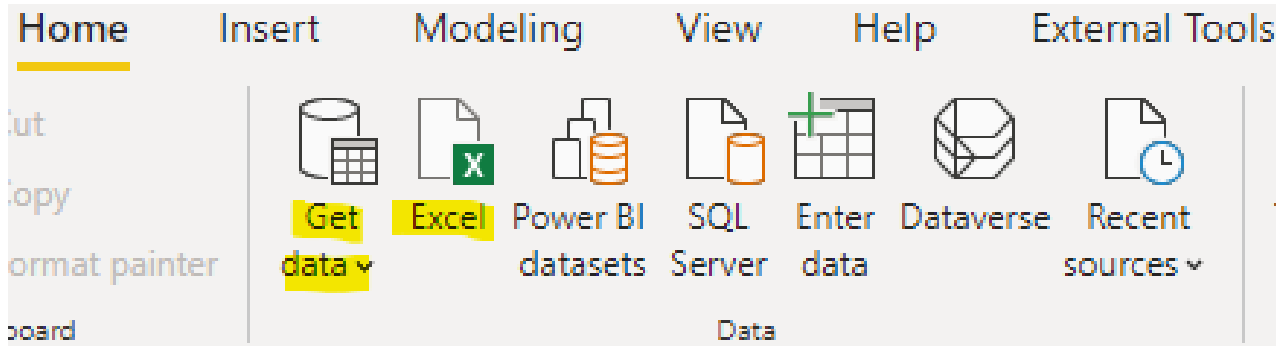
Loading Your Data

You can preview your data before loading it
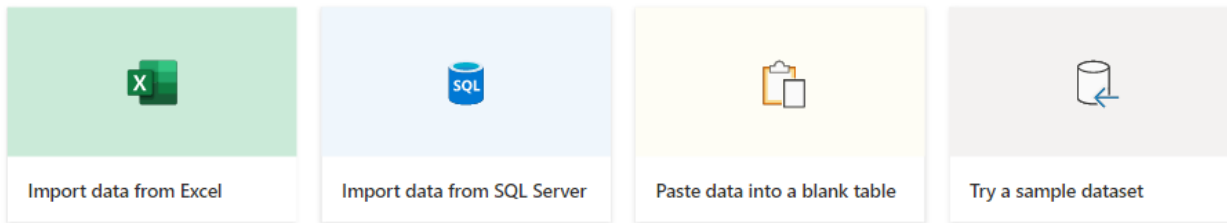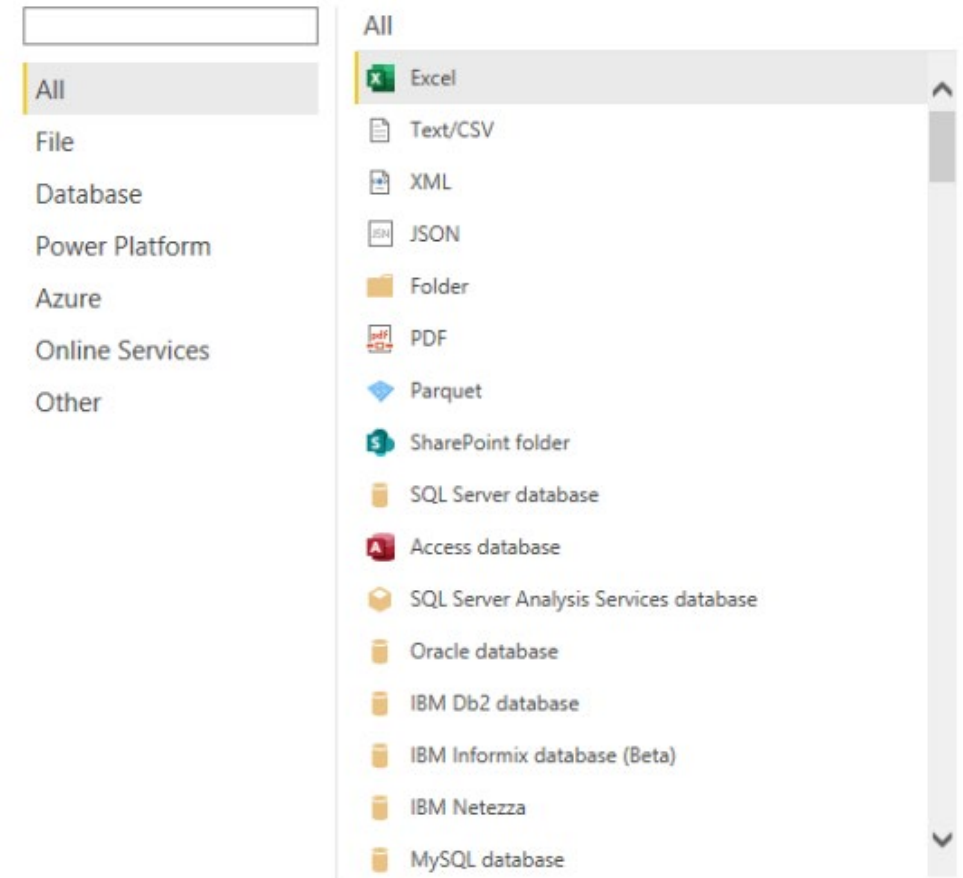
These are the individual spreadsheets from the workbook

Option to transform data prior to loading (can transform after as well)

Creating your data model

Again, we will use Power BI to demonstrate

Review how the data were imported
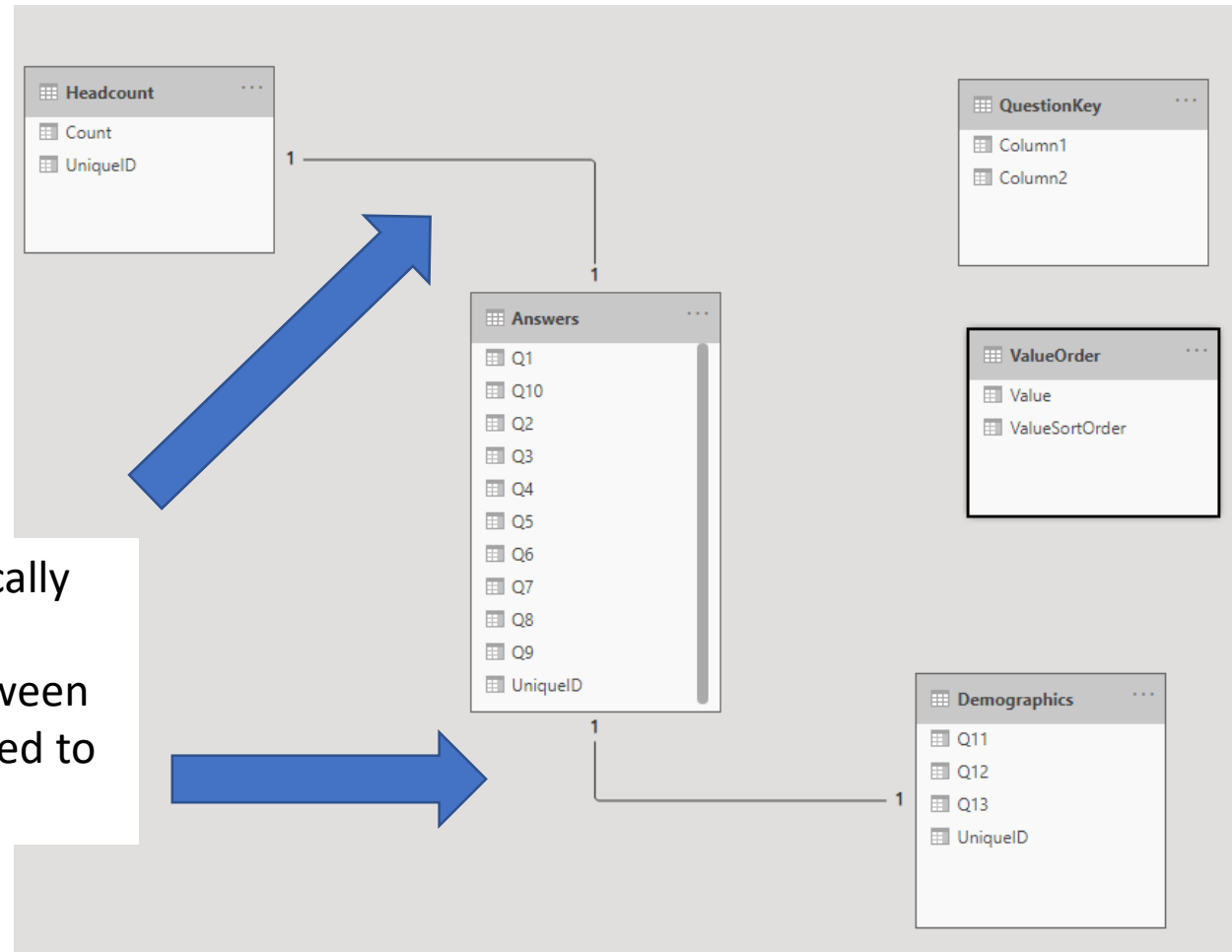
Power BI is automatically assuming there are relationships between the tables, but we need to specify what they are

Review how the data were imported
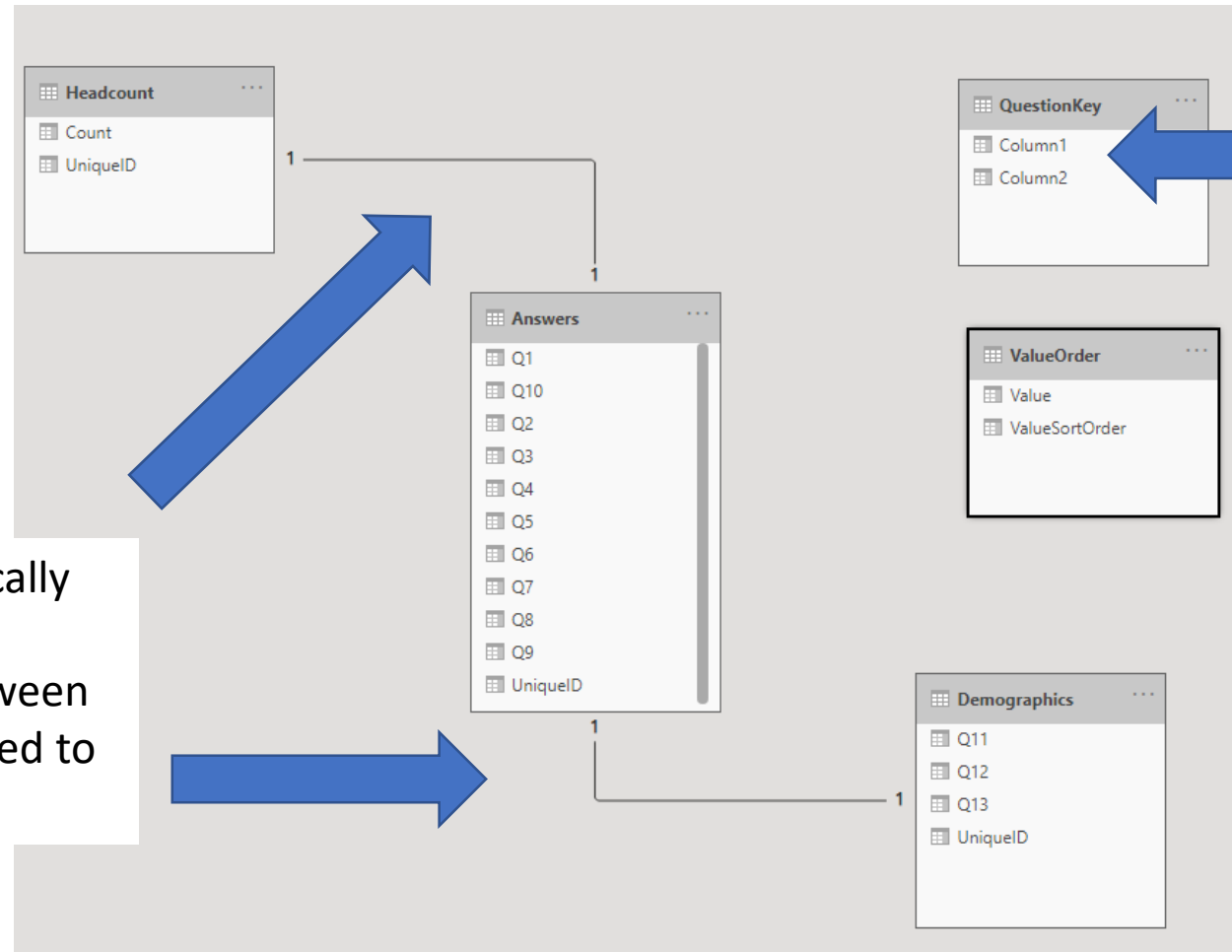
Some column names may not import cleanly, but you can easily fix this

Power BI is automatically assuming there are relationships between the tables, but we need to specify what they are

Promote Column Headers

	ABC Column1	▼	ABC Column2	▼
1	QestionID		QuestionText	
2	Q1		I feel welcome at Faber College.	
3	Q2		I feel valued at Faber College.	
4	Q3		I feel safe at Faber College.	
5	Q4		I think fraternities are an important part of Faber College student life.	
6	Q5		I think togas are appropriate classroom attire.	
7	Q6		I do not appreciate the pop culture references in this survey.	
8	Q7		How often do you skip class to attend fraternity parties?	
9	Q8		How often do you take advantage of Dean Wormer's office hours?	
10	Q9		How would you rate the overall quality of your education Faber Colleg...	
11	Q10		Would you recommend Faber College to others?	
12	Q11		Which of the following best describes your gender identity	
13	Q12		How old are you?	
14	Q13		Which of the following best describes your race or ethnicity?	

File **Home** **Transform** **Add Column** **View**

Group By | Use First Row as Headers ▼ | Transpose | Reverse Rows | Count Rows | Data Type: Text ▼ | Detect Data Type | Rename

Use First Row as Headers

Use Headers as First Row   ⓘ This preview may be u

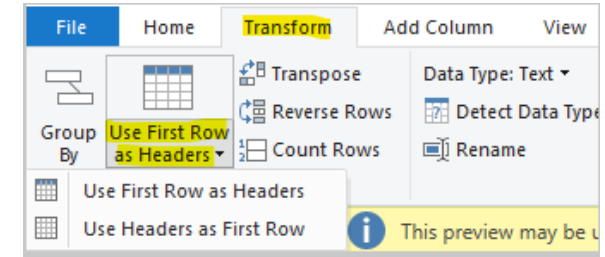|	ABC QestionID	▼	ABC QuestionText	▼
	Q1		I feel welcome at Faber College.	
	Q2		I feel valued at Faber College.	
3	Q3		I feel safe at Faber College.	
4	Q4		I think fraternities are an important part of Faber College student life.	
5	Q5		I think togas are appropriate classroom attire.	
6	Q6		I do not appreciate the pop culture references in this survey.	
7	Q7		How often do you skip class to attend fraternity parties?	
8	Q8		How often do you take advantage of Dean Wormer's office hours?	
9	Q9		How would you rate the overall quality of your education Faber Colleg...	
10	Q10		Would you recommend Faber College to others?	
11	Q11		Which of the following best describes your gender identity	
12	Q12		How old are you?	
13	Q13		Which of the following best describes your race or ethnicity?	
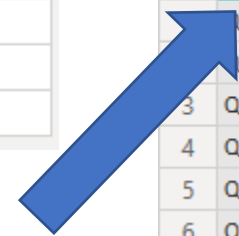
Make sure to fix any typos :)

To do this in Power BI, double click the column title.

Unpivot

Open PowerQuery
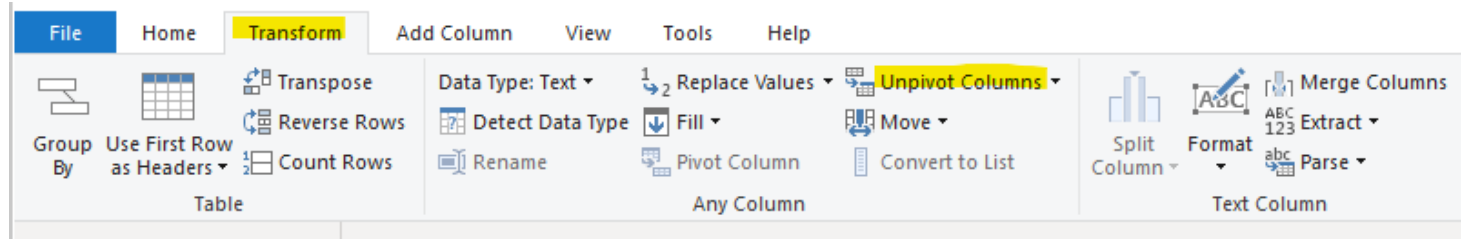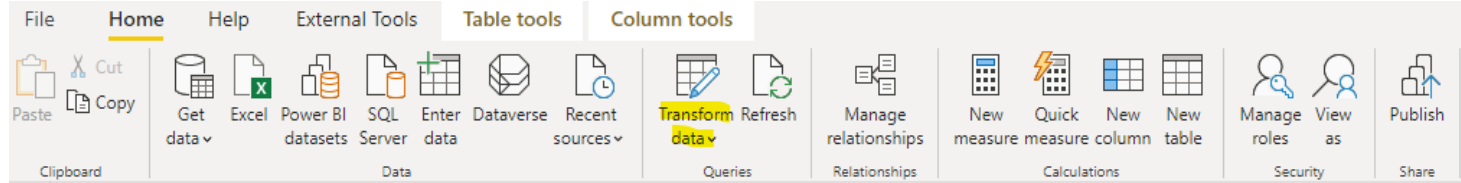
You'll want to rename this to QuestionID

Wide Format

Long Format

Relationships

One-to-one (1:1) - the record only connects to one other record

One-to-many (1:*) - one record can connect to many records

ex. A sports team: one team, multiple players.

Many-to-many (*:*) - each record can connect to multiple records

ex. Students to classes for a semester. **Each class** has **multiple students** enrolled and **each student** can be enrolled in **multiple classes**

Establish Relationships
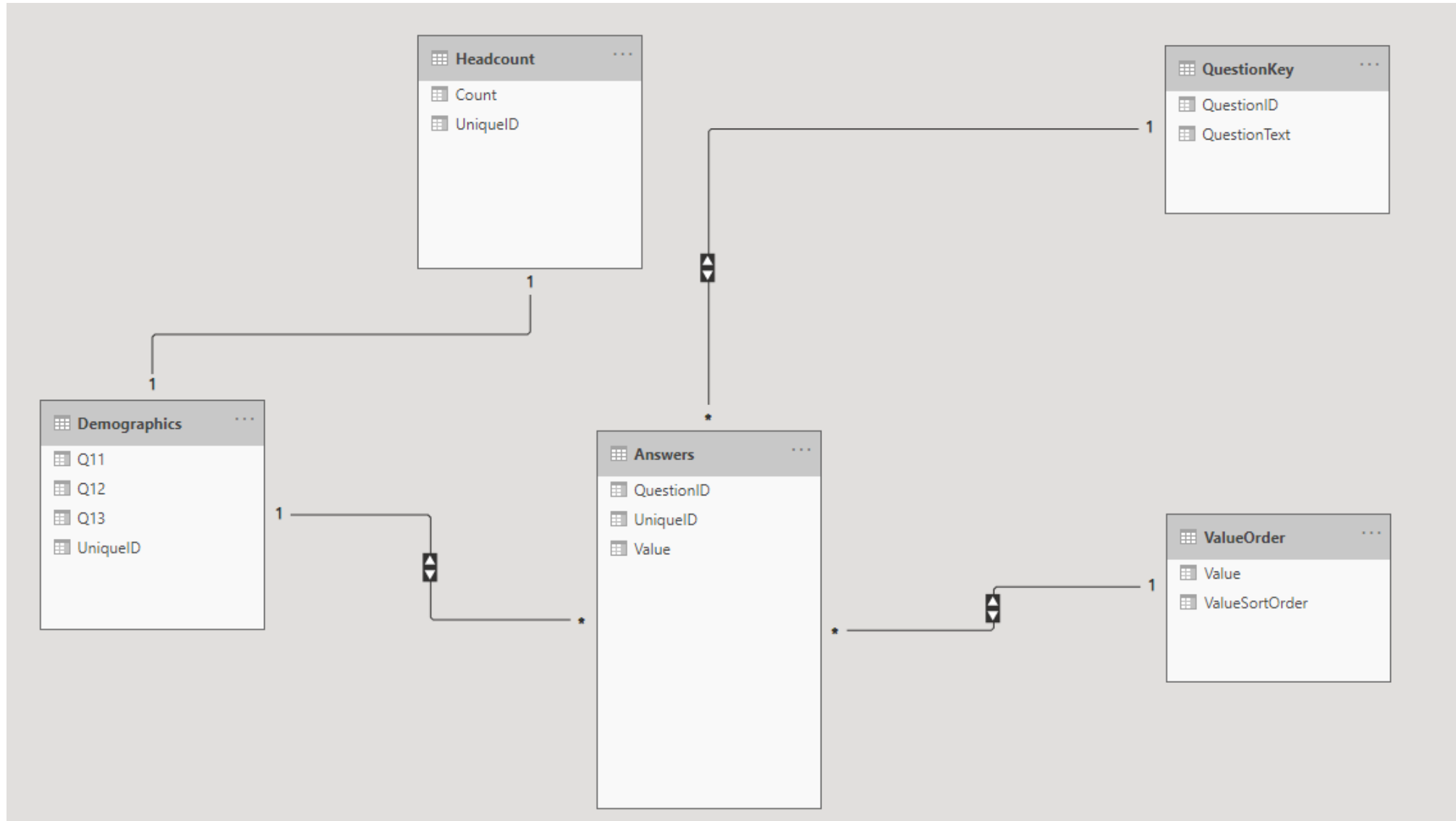
Establish Relationships

Establish Relationships

Establish Relationships

Manage & Edit Relationships

Manage relationships

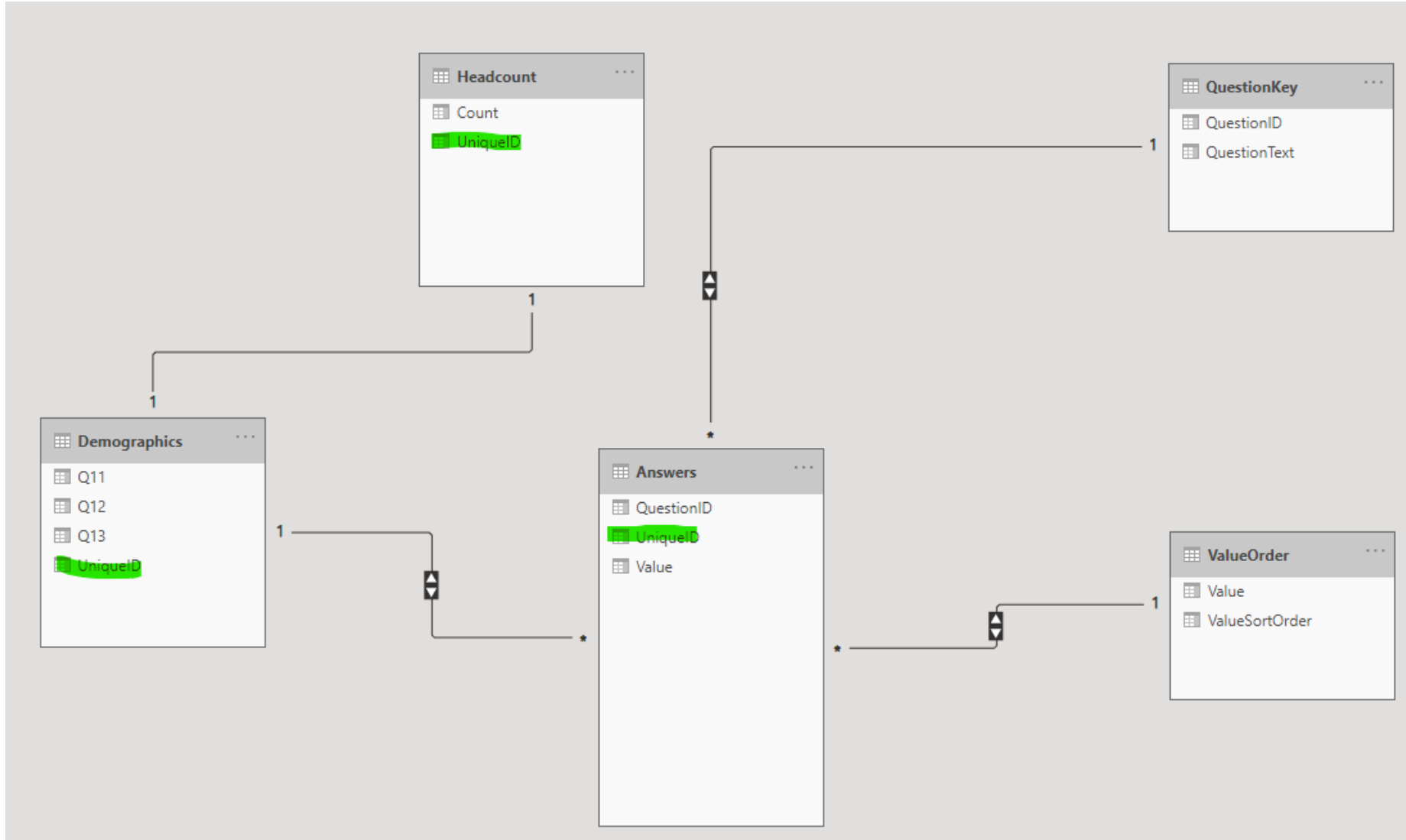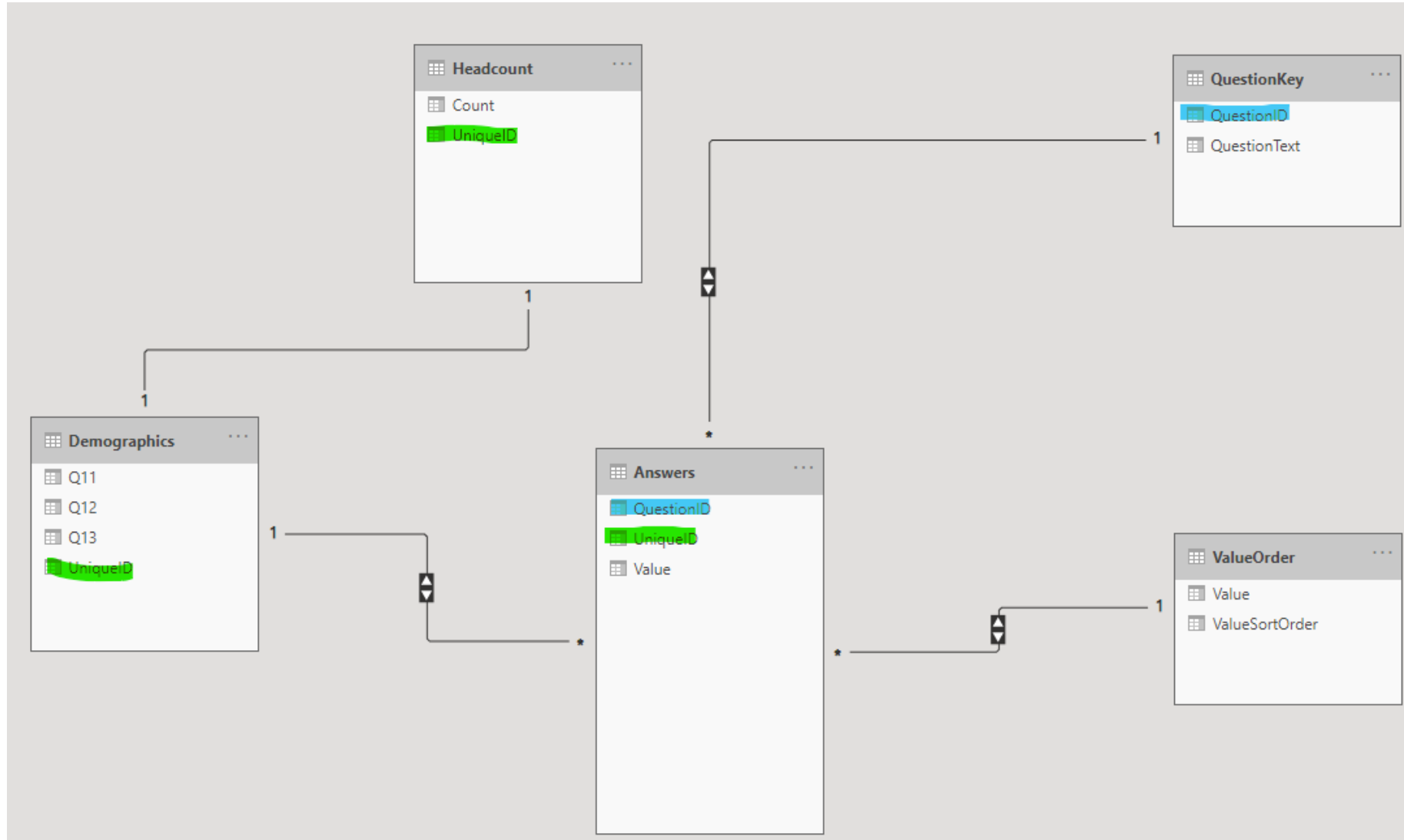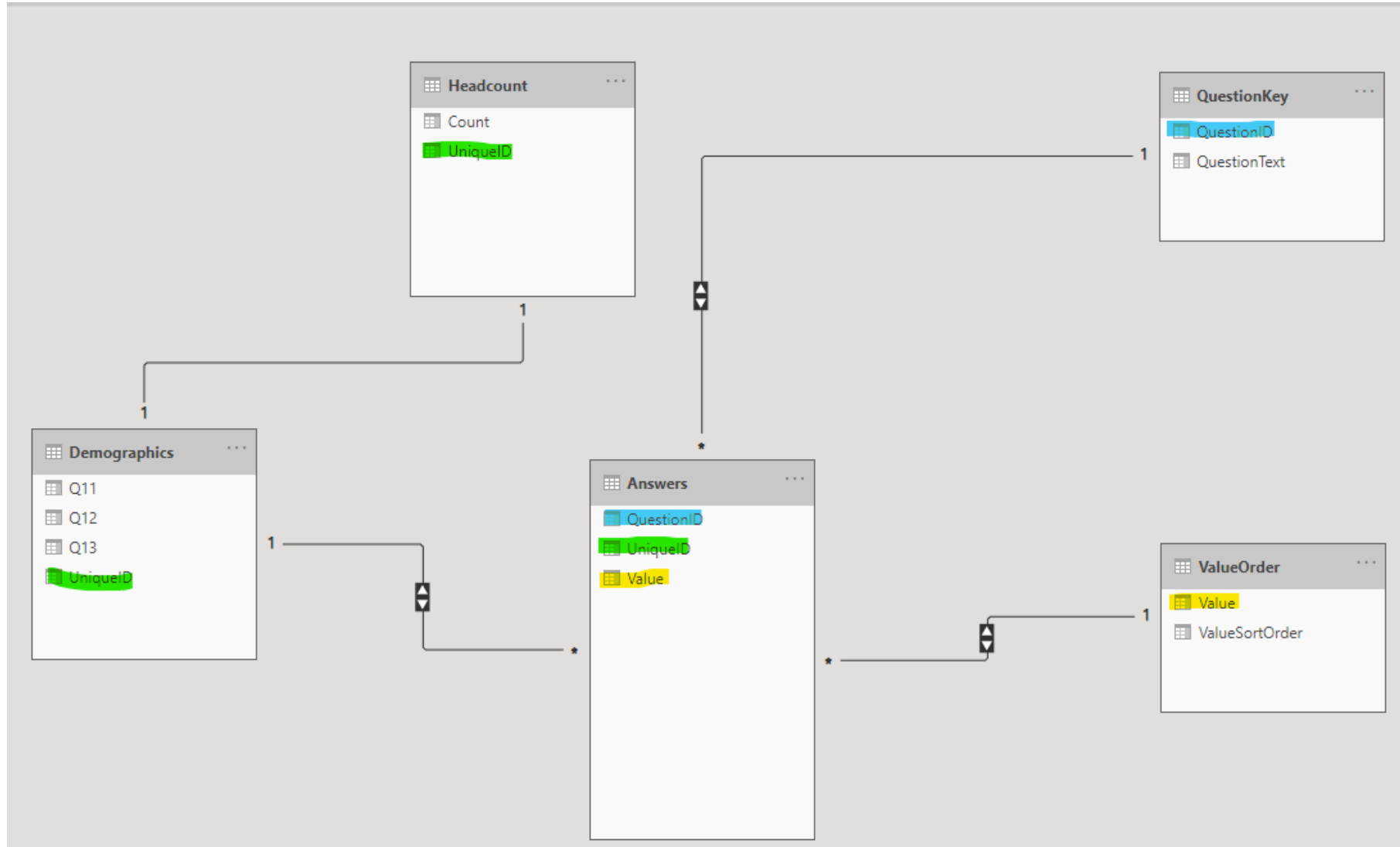| Active	From: Table (Column)	To: Table (Column)
☑	Answers (QuestionID)	QuestionKey (QuestionID)
☑	Answers (UniqueID)	Demographics (UniqueID)
☑	Answers (Value)	ValueOrder (Value)
☑	Demographics (UniqueID)	Headcount (UniqueID)

New...   Autodetect...   Edit...   Delete

Edit relationship

Select tables and columns that are related.

Answers ▾

UniqueID	QuestionID	Value
53	Q1	Strongly agree
54	Q1	Strongly agree
55	Q1	Strongly agree

QuestionKey ▾

QuestionID	QuestionText
Q1	I feel welcome at Faber College.
Q2	I feel valued at Faber College.
Q3	I feel safe at Faber College.

Cardinality   Cross filter direction

Many to one (*:1) ▾   Both ▾

☑ Make this relationship active   ☐ Apply security filter in both directions

☐ Assume referential integrity

OK   Cancel
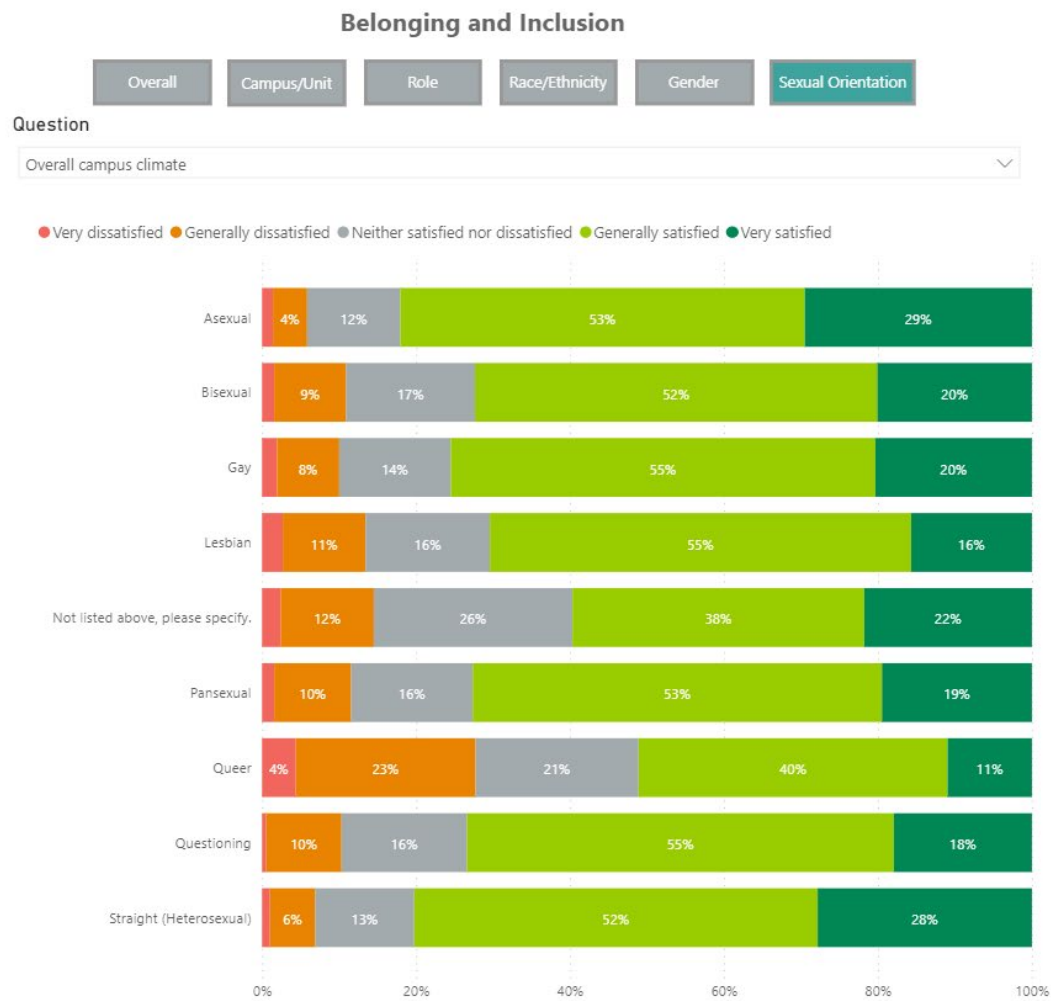
With a solid data model for your foundation, you are ready to start building

The image to the left is from Penn State's Community Survey dashboard, https://opair.psu.edu/community-survey/dashboards/

Just a sample of the many resources available to help you on your data visualization journey

- Microsoft Power BI
  - Guy in a Cube, https://www.youtube.com/watch?v=B_bAKrXAUh0
  - Microsoft Learning Paths, https://docs.microsoft.com/en-us/learn/powerplatform/power-bi?WT.mc_id=powerbi_landingpage-docs-link
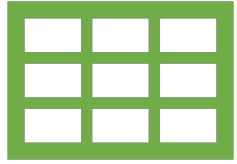 - Microsoft Power BI YouTube, https://www.youtube.com/playlist?list=PL1N57mwBHtN0JFoKSR0n-tBkUJHeMP2cP&app=desktop
  - Power BI Blog, https://powerbi.microsoft.com/en-us/blog/
- Tableau
  - Tableau Training videos, https://www.tableau.com/learn/training/20202
- LinkedIn Learning, Coursera, Udemy, Udacity, Edx

In recap, we hope that you can now:

Describe flat files and their limitations
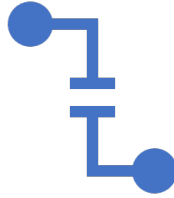
Sketch out a basic relational data model

List the key steps to build a data model

Find additional resources for self-guided learning

...so that you can get on with the fun part – visualizing your data!

Thank you!

- Betty Harper, bharper@psu.edu
- Shannon Lee, slee@psu.edu
- Carly Sunseri, cjt23@psu.edu