

The background of the slide is a grayscale image of a circuit board. It features a complex network of black lines representing traces, with several large black circular pads or vias. Faint, concentric circular patterns are visible in the background, suggesting a microscopic view of the board's surface.

Intro to Challenge 1

Data Sets

what you need to know

Challenge 1: Data Sets

Status: Challenge Complete

What to do next: All Done!

1. Get the Skills you Need

Study the practice questions in these levels. Mark them "Ready for Challenge" when you are done.

Python skill level(s) you need



Level 1: What's Programming?



Level 2: In and out of Python

Data Science skill level(s) you need



Level 1: What's Data Science?



Level 2: What's a data set?

Statistics skill level(s) you need



Level 1: What's Statistics?



Level 2: What's Data?

What you do

1. Download and open a data set
2. Talk about the social context of the data
3. Talk about the data itself
4. Do a data analysis to find answers

Key Data Science Points

- What is in data science that is not in stats and python?
- What is a data context?
- What is a data set?

Key Stats Points

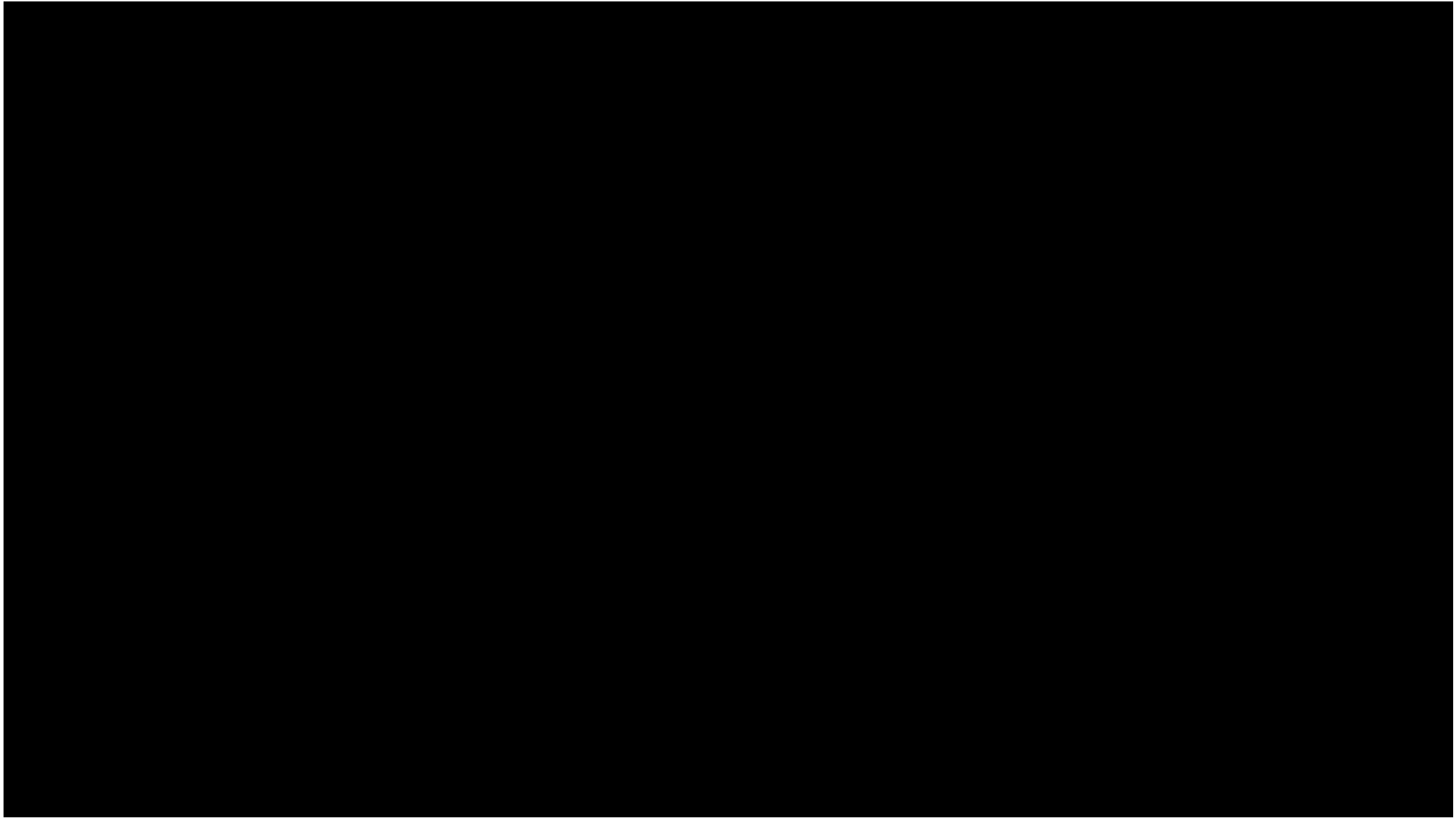
- What is Data?
- What's a data type?
- What is a table
- Row
- Column

The screenshot shows an Excel spreadsheet titled "mayorsfoodcourt_filtered.csv - Excel". The ribbon is set to "Home", and the "Clipboard" group is active. The formula bar shows "businessName" in cell A1. The data table has the following columns: businessName, DESCRPT, RESULTDTM, Violation, ViolLevel, and ViolD. The rows contain data for "100 Percent Delicia Food" with various violations related to "Eating & Drinking".

	A	B	C	D	E	F
1	businessName	DESCRPT	RESULTDTM	Violation	ViolLevel	ViolD
2	100 Percent Delicia Food	Eating & Drinking	2/15/2013 12:19	15-4-202.1*		Non-F
3	100 Percent Delicia Food	Eating & Drinking	2/15/2013 12:19	16-4-501.1*		Dishw
4	100 Percent Delicia Food	Eating & Drinking	2/15/2013 12:19	16-4-501.1*		Dishw
5	100 Percent Delicia Food	Eating & Drinking	2/15/2013 12:19	29-5-201/i*		Instal
6	100 Percent Delicia Food	Eating & Drinking	2/15/2013 12:19	31-5-204/i***		Locat
7	100 Percent Delicia Food	Eating & Drinking	2/15/2013 12:19	37-6-501.1*		Impro
8	100 Percent Delicia Food	Eating & Drinking	3/21/2013 14:48	16-4-501.1*		Dishw
9	100 Percent Delicia Food	Eating & Drinking	10/24/2013 11:18	05-4-302.1*		Food
10	100 Percent Delicia Food	Eating & Drinking	10/24/2013 11:18	14-4-202.1*		Food
11	100 Percent Delicia Food	Eating & Drinking	10/24/2013 11:18	16-4-501.1*		Dishw
12	100 Percent Delicia Food	Eating & Drinking	10/24/2013 11:18	16-4-501.1*		Dishw
13	100 Percent Delicia Food	Eating & Drinking	10/24/2013 11:18	32-6-301.1*		Hand
14	100 Percent Delicia Food	Eating & Drinking	10/24/2013 11:18	37-6-501.1*		Impro
15	100 Percent Delicia Food	Eating & Drinking	6/19/2014 14:58	29-5-201/i*		Instal
16	100 Percent Delicia Food	Eating & Drinking	6/19/2014 14:58	36-6-501.1*		Impro

Key Python Points

- `import pandas as pd`
- `dataFrame = pd.read_csv(", sep="")`
- `print()`
- `dataFrame.head()`



The background of the slide is a grayscale image of a circuit board. It features a complex network of black lines representing traces, with several large black circular pads or vias. Faint, concentric circular patterns are visible in the background, suggesting a microscopic view of the board's surface.

Intro to Challenge 2

Describing Data

what you need to know

Challenge 2: Describing Data

Status: Challenge Complete

What to do next: All Done!

1. Get the Skills you Need

Study the practice questions in these levels. Mark them "Ready for Challenge" when you are done.

Python skill level(s) you need



Level 3: Making Calculations in Python



Level 4: Python Data Structures

Data Science skill level(s) you need



Level 3: Describing Data: Google Trends

Statistics skill level(s) you need



Level 4: Working with Graphs and Charts



Level 3: Describing Data

What you do

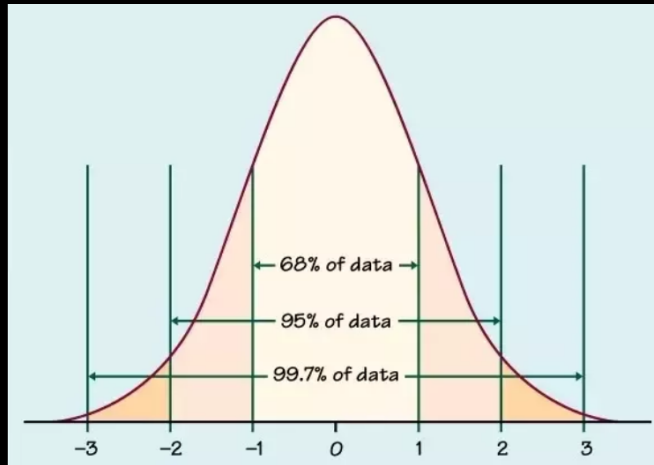
1. Download and open a data set
2. Talk about the social context of the data
3. Talk about the data itself
4. Do a data analysis to find answers

Key Data Science Points

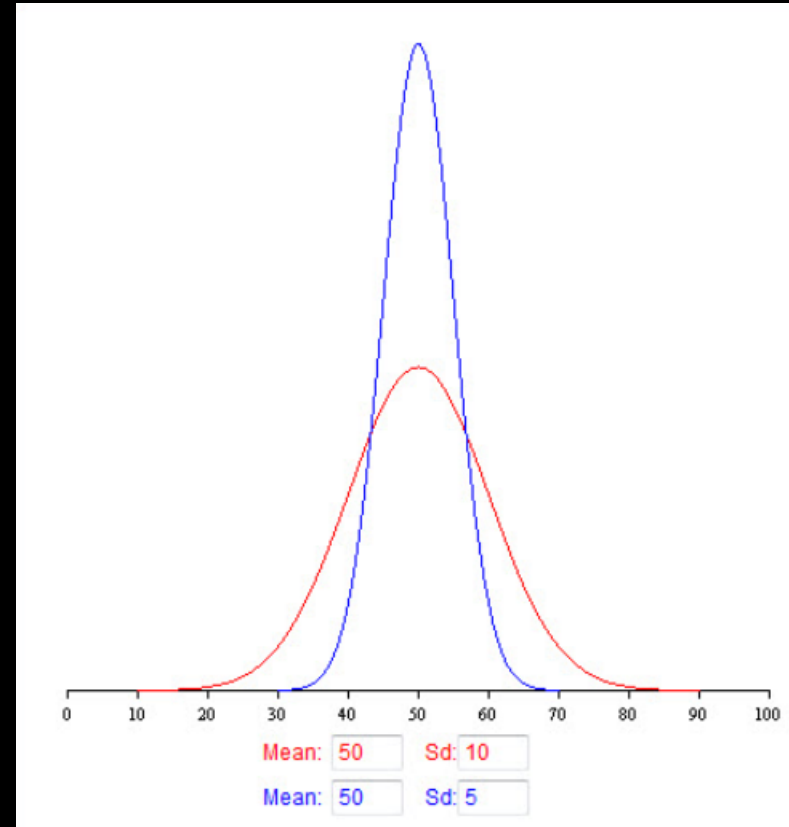
- Bias
- Normalized
- Anonymized
- Missing values
- Google Trends

Key Stats Points

- Central tendency
- Mean
- Standard deviation



[Quora](#)



[OnlineStatBook.com](#)

Key Python Points

- `df = pd.read_csv("")`
- `meanAsNumber = df[""].mean()`
- `meanAsMoney = "${{}}".format(meanAsNumber)`

