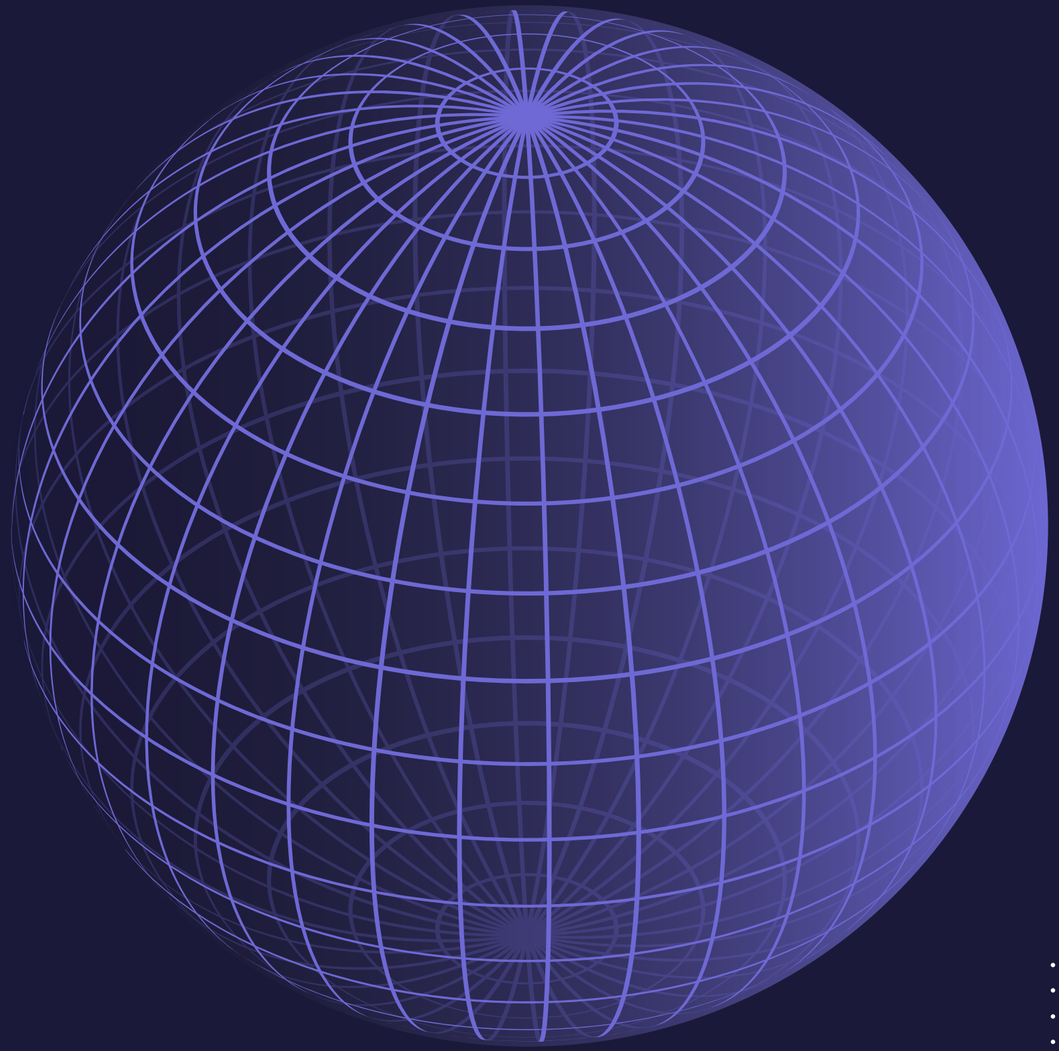
001

# GETTING STARTED

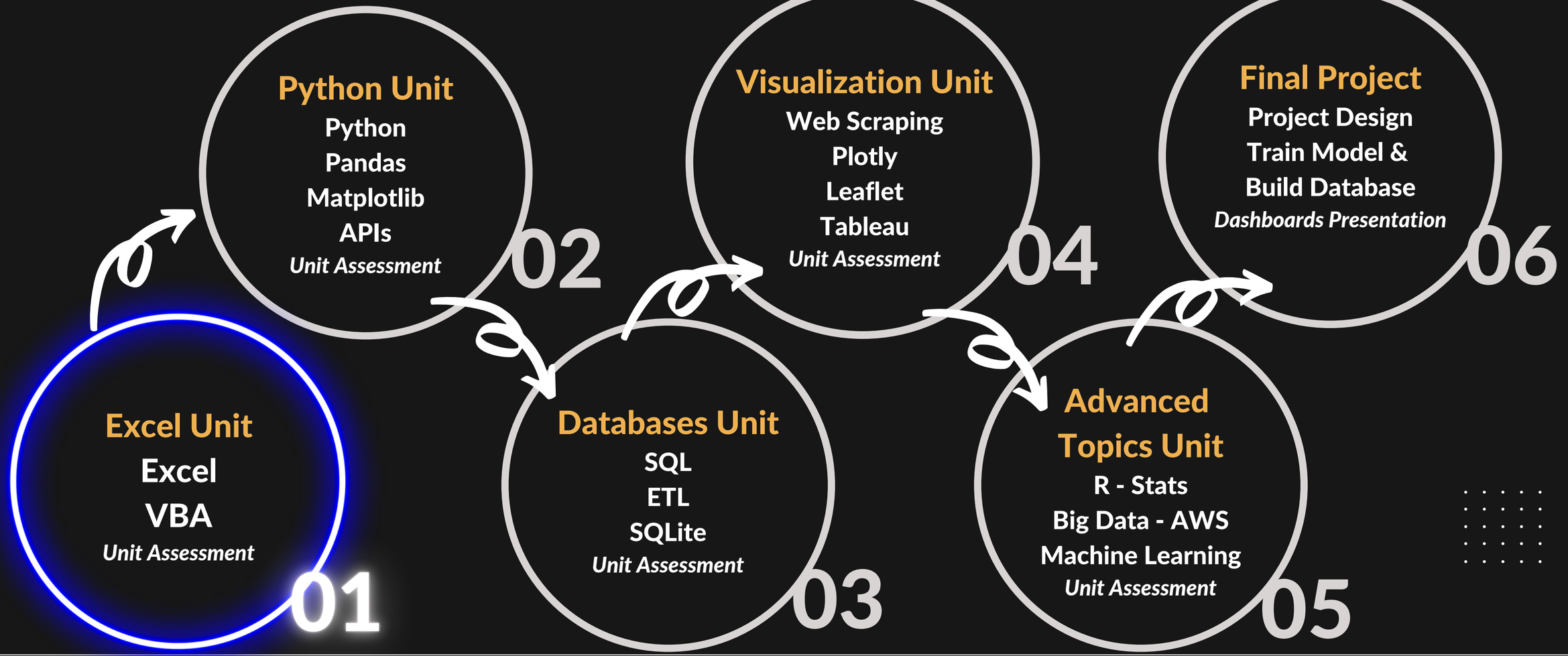


DATA BOOTCAMP
UNIVERSITY OF KANSAS

UNIVERISTY OF KANSAS

# THE BIG PICTURE





# BOOT CAMP POINTERS

## As you work through this module, remember the following:

**01**

Your Bootcamp Spot material is connected to the things we will do during class.
It's all part of the journey!

**02**

Your coursework this week will prepare you with all the skills that you need to succeed on your Challenge assignment!

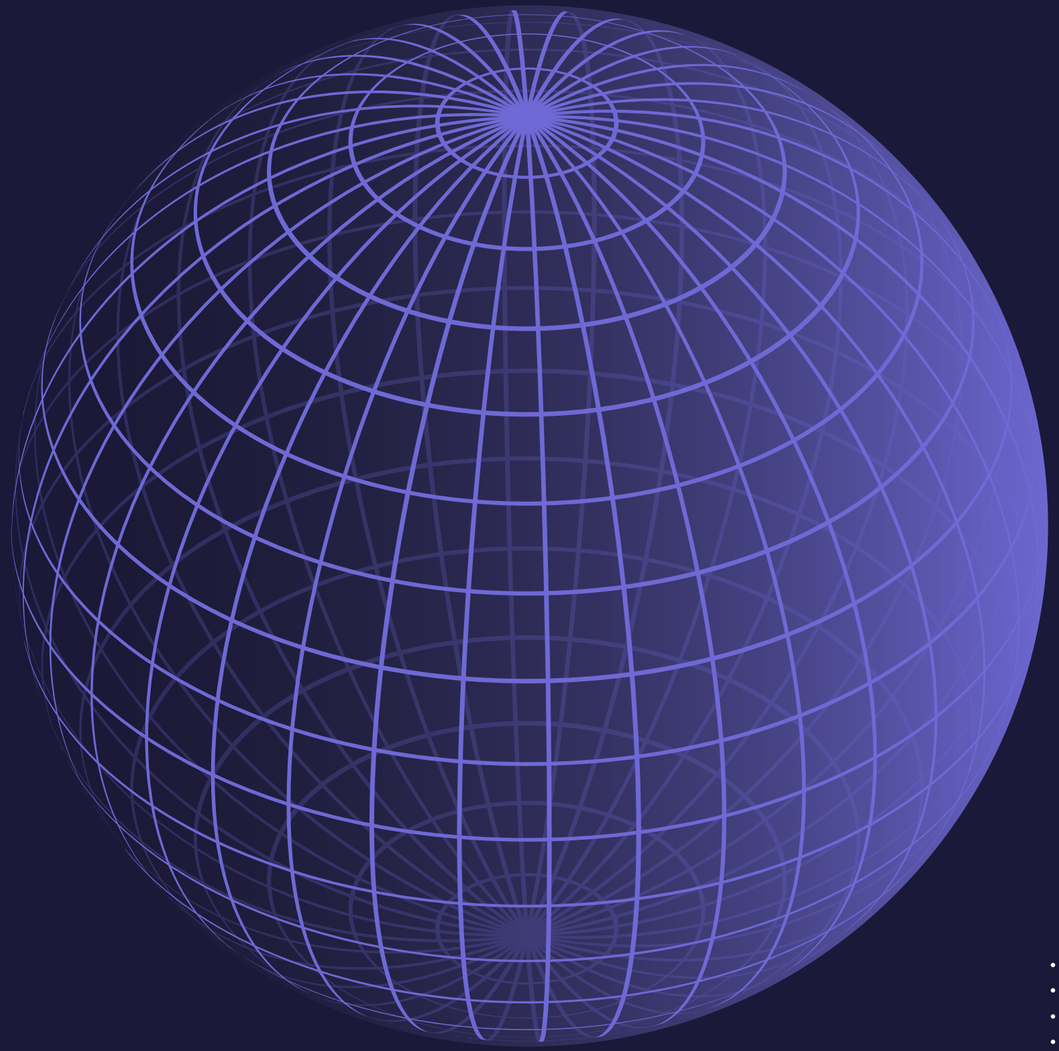
**03**

Be proactive in using Office Hours to get help with any installation issues.
We're here to help!

004

# UNIT 1: EXCEL

DATA BOOTCAMP
UNIVERSITY OF KANSAS



ZEND OF DATA | UNIVERISTY OF KANSAS

# THIS WEEK: EXCEL



## By the end of this week, you'll know how to:

- Import data into Excel
- Apply filters, conditional formatting, and formulas to data
- Create and interpret charts and pivot tables in Excel
- Calculate summary statistics
- Characterize data to identify outliers in datasets
- Visualize the distribution of data using box plots

# THIS WEEK'S CHALLENGE

## Kickstarter Challenge

**Analyze a dataset consisting of 4,000 crowdfunding projects to discover hidden trends.**

Using pivot tables and functions to filter data...

...create charts that demonstrate an analysis of data sets to visualize business outcomes based on launch dates and goals.

- **Deliverable 1**: Outcomes Based on Launch Date Chart
- **Deliverable 2:** Outcomes Based on Goals Chart
- **Deliverable 3:** A written analysis of the results (README.md)

# CAREER CONNECTIONS

How will you use this module's content in your career?

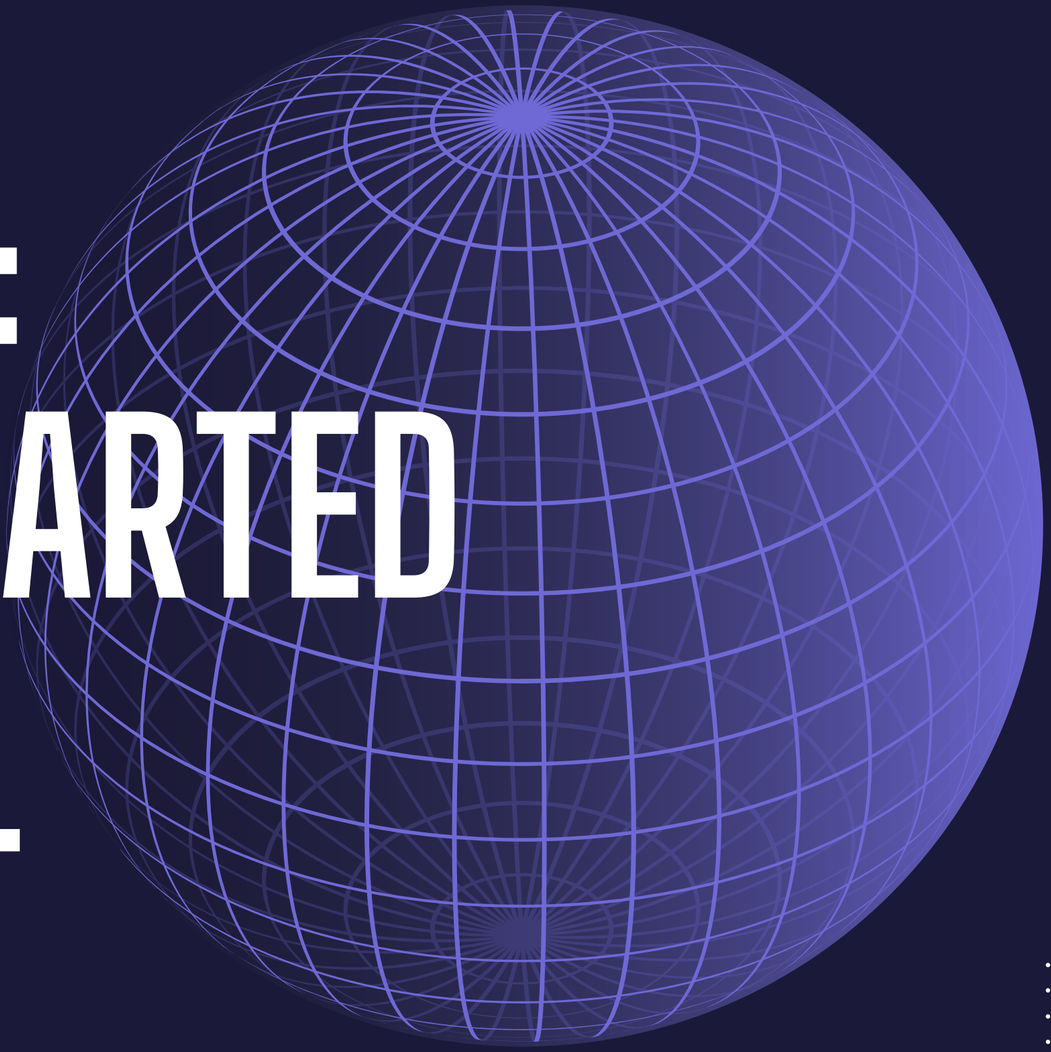
# QUICK TIP FOR SUCCESS

There are hundreds of Excel functions.

You'll have to look some of them up.

Consider this your first opportunity to dive into some documentation!

# MODULE 1.1: GETTING STARTED WITH EXCEL



DATA BOOTCAMP
UNIVERSITY OF KANSAS

# TODAY'S AGENDA



## By completing today’s activities, you’ll learn the following skills:

- Calculating Averages
- Pivot Tables & Pivot Charts
- Lookups



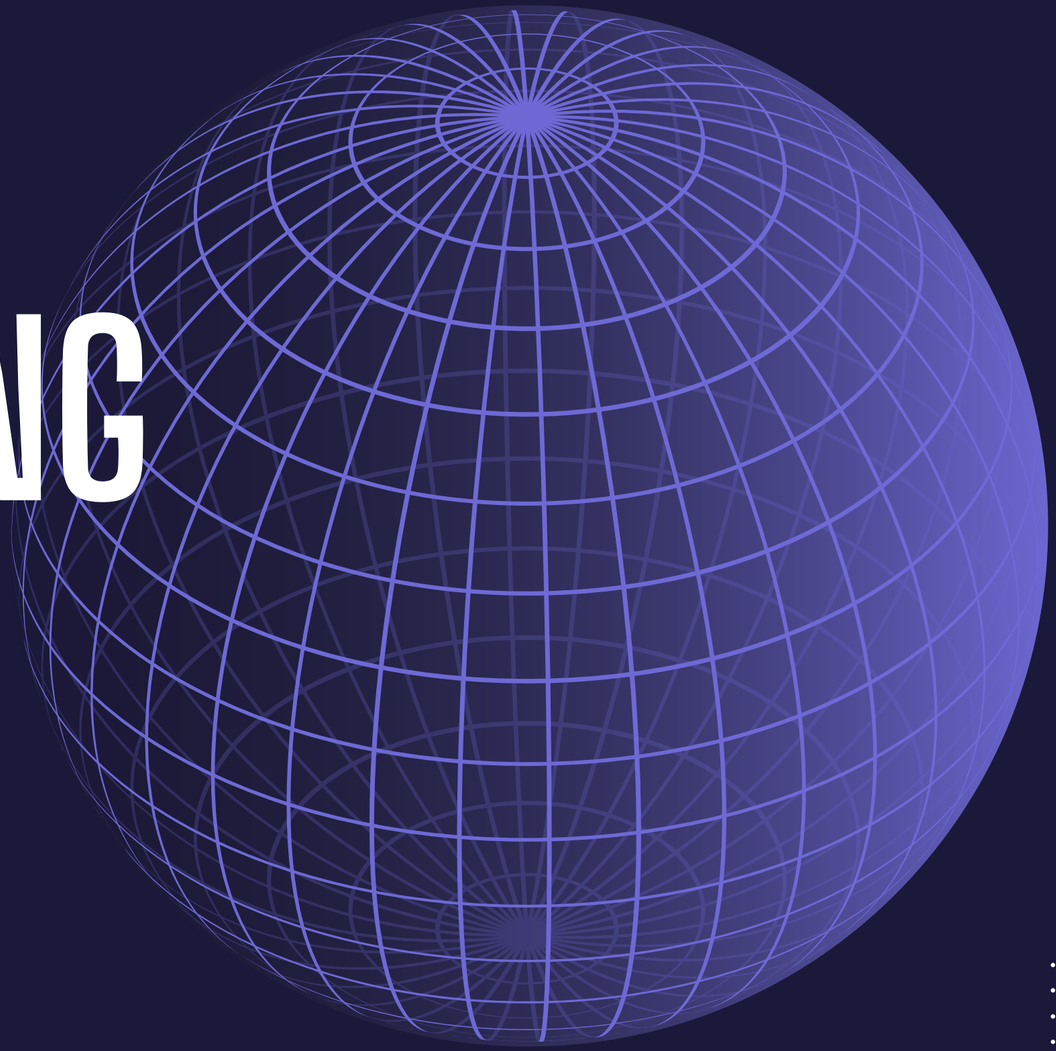
Make sure you’ve downloaded any relevant class files!

# CHECK IN

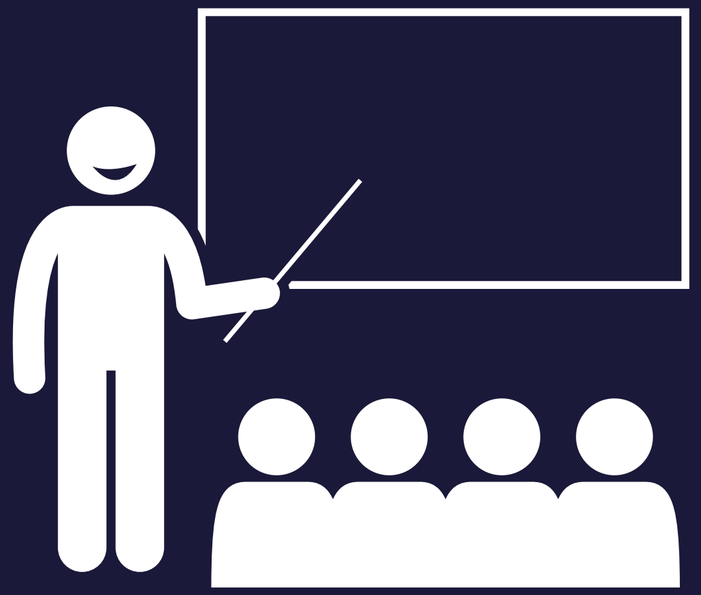
UNIVERSITY OF KANSAS

010

# CALCULATING AVERAGES



|UNIVERISTY OF KANSAS



# INSTRUCTOR DEMONSTRATION

## Reviewing Averages

UNIVERSITY OF KANSAS

# ACTIVITY: GRADEBOOK

Create a formula that calculates the final grade for a student based on their previous exams and papers.

**Suggested Time: 15 minutes**

# ACTIVITY: GRADEBOOK

**Suggested Time: 15 minutes**

## To do:

Create a formula that calculates the final grade for a student based upon their previous exams and assignments.

- What is a student's final grade?
- Did the student pass the class?
- What is the student's letter grade?

## When making this calculation:

- **Consider every assignment and exam to be equal in weight; each should comprise one-fourth of the overall grade.**
- **Round the result to the nearest integer.**
- **Using conditionals, create a formula that returns PASS if a student's final grade is greater than or equal to 60. If a student's final grade is below 60, the formula should return FAIL.**

## Bonus:

**Create a nested IF() formula that returns a letter grade based on a student's final grade.**

- **Greater than or equal to 90 = A**
- **Greater than or equal to 80 and less than 90 = B**
- **Greater than or equal to 70 and less than 80 = C**
- **Greater than or equal to 60 and less than 70 = D**
- **Anything less than 60 = F**

# CHECK IN

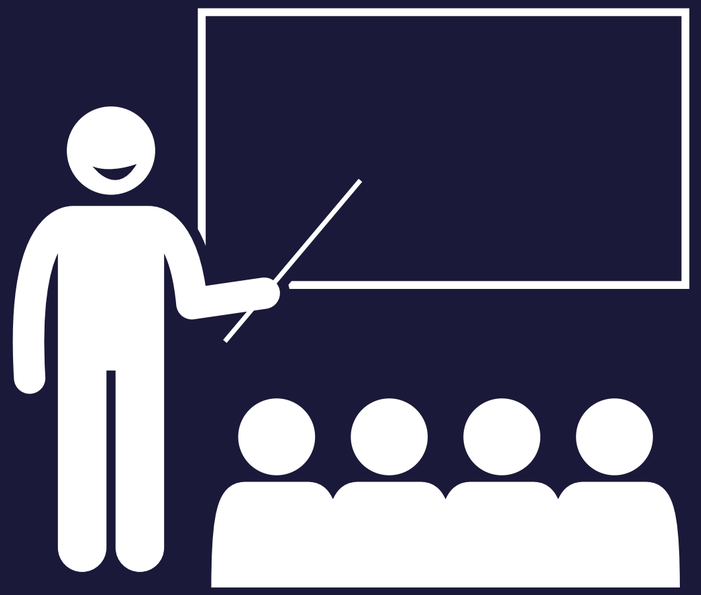
Let's Review

UNIVERSITY OF KANSAS

010

# PIVOT TABLES & PIVOT CHARTS

|UNIVERISTY OF KANSAS



# INSTRUCTOR DEMONSTRATION

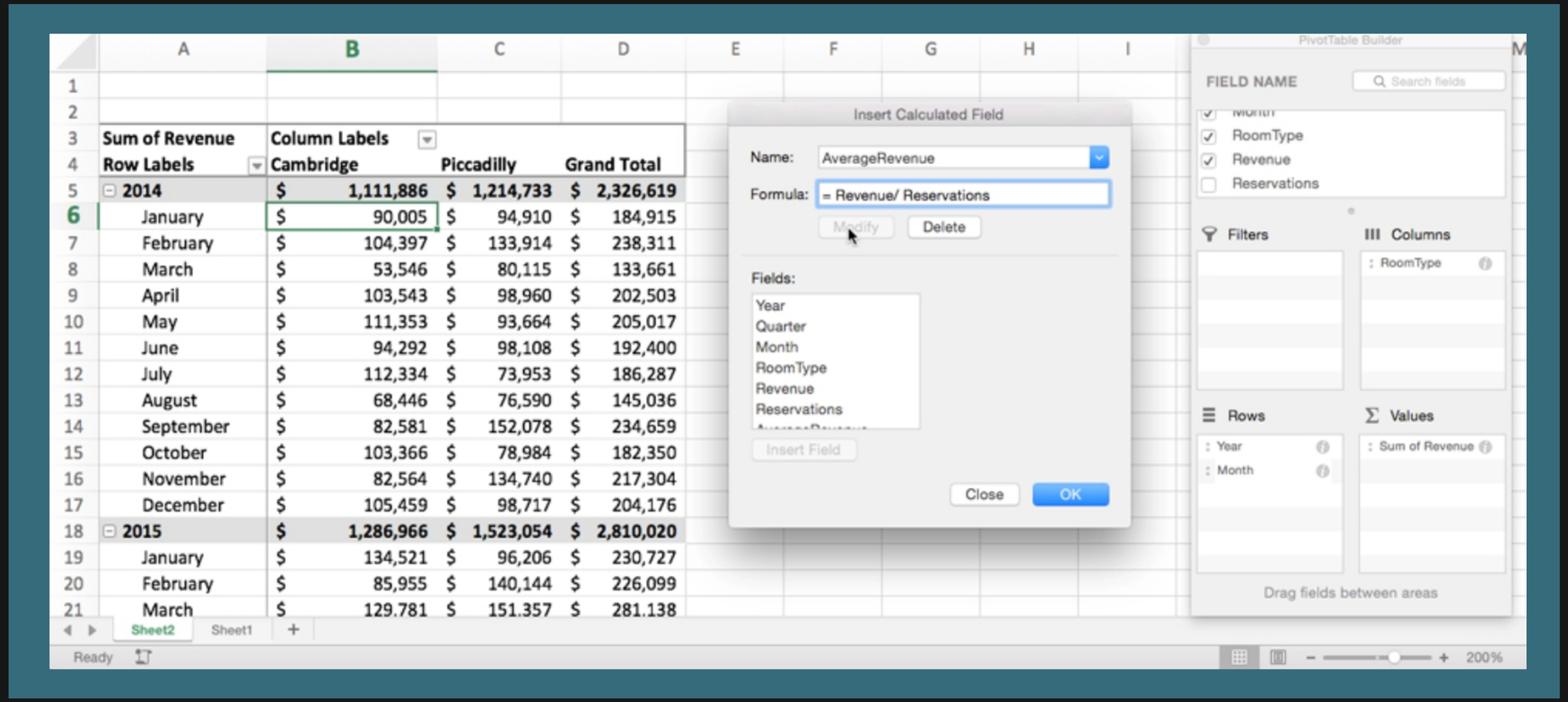
Pivot Tables & Pivot Chart

UNIVERSITY OF KANSAS

# GET PIVOT WITH IT



**Pivot tables are one of the most important data visualisation concepts to master in this class (don't worry, they are a cinch to deal with).**




| Sum of Revenue | Column Labels | | |
|---|---|---|---|
| Row Labels | Cambridge | Piccadilly | Grand Total |
| 2014 | $ 1,111,886 | $ 1,214,733 | $ 2,326,619 |
| January | $ 90,005 | $ 94,910 | $ 184,915 |
| February | $ 104,397 | $ 133,914 | $ 238,311 |
| March | $ 53,546 | $ 80,115 | $ 133,661 |
| April | $ 103,543 | $ 98,960 | $ 202,503 |
| May | $ 111,353 | $ 93,664 | $ 205,017 |
| June | $ 94,292 | $ 98,108 | $ 192,400 |
| July | $ 112,334 | $ 73,953 | $ 186,287 |
| August | $ 68,446 | $ 76,590 | $ 145,036 |
| September | $ 82,581 | $ 152,078 | $ 234,659 |
| October | $ 103,366 | $ 78,984 | $ 182,350 |
| November | $ 82,564 | $ 134,740 | $ 217,304 |
| December | $ 105,459 | $ 98,717 | $ 204,176 |
| 2015 | $ 1,286,966 | $ 1,523,054 | $ 2,810,020 |
| January | $ 134,521 | $ 96,206 | $ 230,727 |
| February | $ 85,955 | $ 140,144 | $ 226,099 |
| March | $ 129,781 | $ 151,357 | $ 281,138 |

# GET PIVOT WITH IT



**Essentially, a pivot table is a summative analytic tool that allows us to perform aggregate functions that allow any combination of fields. The term pivot table comes from the fact that we are pivoting along a data axis.**

| Seller | Quantity Sold | Date |
|---|---|---|
| Joseph | $42.50 | 1/1/17 |
| Jacob | $65.00 | 1/3/17 |
| Jacob | $5.25 | 1/6/17 |
| Joseph | $125.00 | 1/6/17 |
| Jacob | $3.50 | 1/7/17 |
| Matt | $32.00 | 1/9/17 |

| Seller | Total Sold |
|---|---|
| Joseph | $167.50 |
| Jacob | $73.75 |
| Matt | $32.00 |

# WORD TO THE WISE: KEEP IT FLAT!



**Modern business intelligence (BI) tools like Tableau, Sisense, and Salesforce work best if data is stored in flat CSVs—meaning column headers represent fields (vertically) on the spreadsheet. This is largely because all of these technologies heavily utilise pivot tables as a tool for their visualisations. Don't try to confuse this simplicity. "Spreadsheet magic" is a nightmare to analyze.**

| B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|
| DateTime | Week # | Section? | Pace | Academic Support | Self-Mastery | Instructor Er |
| 2016-09-11T04:00:00.000Z | 18 | RCB0503FSF - CCC | 3 | 5 | 5 | 4 |
| 2016-09-11T05:00:00.000Z | 6 | UT0726FSF | 3 | 5 | 3 | 4 |
| 2016-09-12T04:00:00.000Z | 11 | UCF062016FSF | 4 | 4 | 3 | 5 |
| 2016-09-12T04:00:00.000Z | 23 | UCF0329FSF | 2 | 4 | 5 | 1 |
| 2016-09-12T04:00:00.000Z | 9 | UNC0712FSF | 3 | 4 | 4 | 3 |
| 2016-09-12T04:00:00.000Z | 23 | UCF0328FSF | 4 | 3 | 2 | 3 |
| 2016-09-12T04:00:00.000Z | 6 | RUT0725FSF-NB | 5 | 4 | 4 | 5 |
| 2016-09-12T04:00:00.000Z | 6 | RUT0725FSF-NB | 5 | 5 | 4 | 5 |
| 2016-09-12T04:00:00.000Z | 6 | RUT0725FSF-NB | 2 | 4 | 4 | 4 |
| 2016-09-12T04:00:00.000Z | 11 | UCF062016FSF | 4 | 5 | 4 | 5 |
| 2016-09-12T04:00:00.000Z | 13 | UCF061416FSF | 4 | 5 | 1 | 5 |

# ACTIVITY:
# TOP SONGS PIVOT TABLE

In this activity, you will use a 5000-row spreadsheet containing data on the top 5000 songs from 1901 on. Using pivot tables, you will uncover which artists have the most songs in the top 5000, what the songs are, and what year they were released.

**Suggested Time:**
**15 minutes**

UNIVERSITY OF KANSAS

# TOP SONGS PIVOT TABLE INSTRUCTIONS



- Select all of the data in your worksheet and create a new pivot table.
- Make a pivot table that can be filtered by year and contains two rows: *Artist* and *Name*.
- All of an artist's songs should be listed below their name.
- Update your pivot table to contain values for:
  - How many songs an artist has in the top 5000
  - The sum of the final_score of their songs.
- Sort your pivot table by descending sum of the final_score.

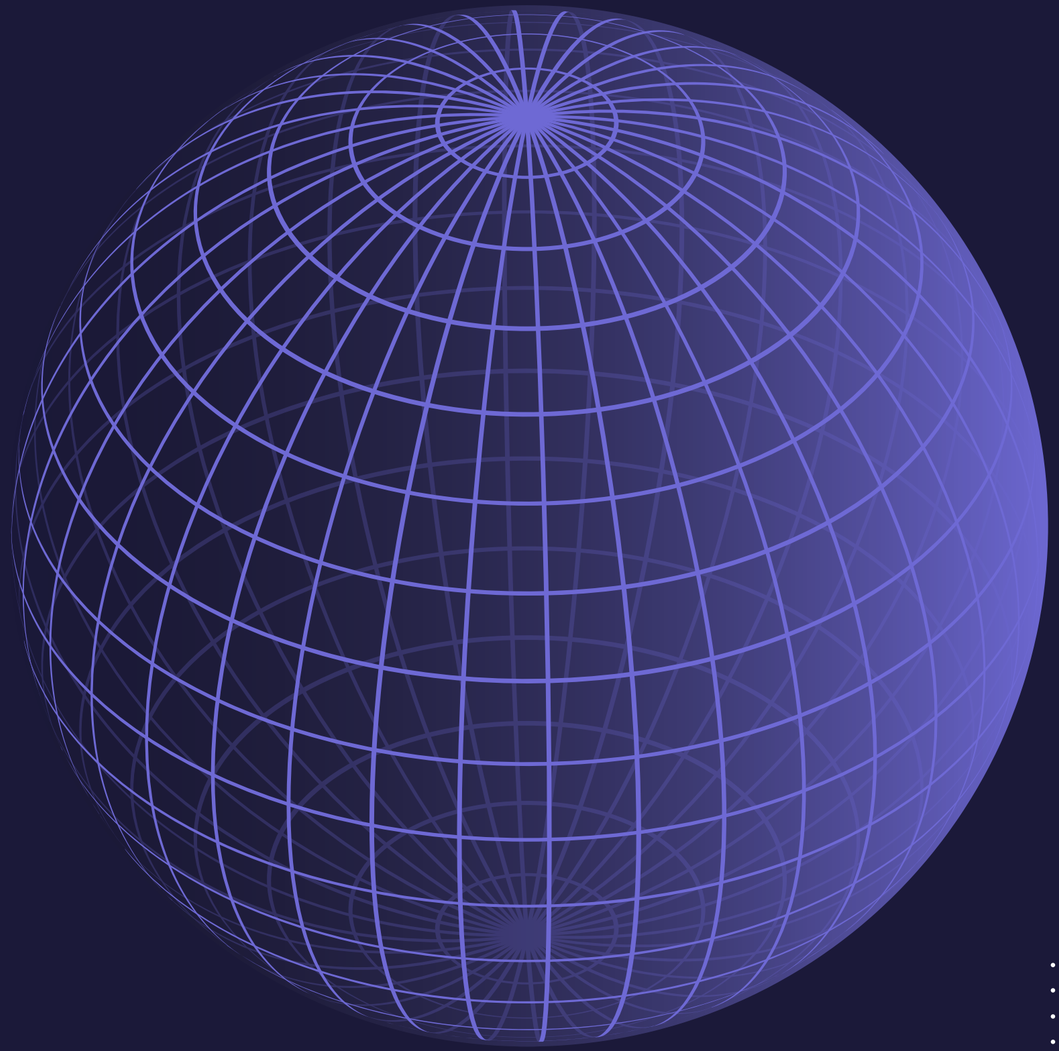
**Suggested Time: 15 minutes**

# CHECK IN

Let's Review

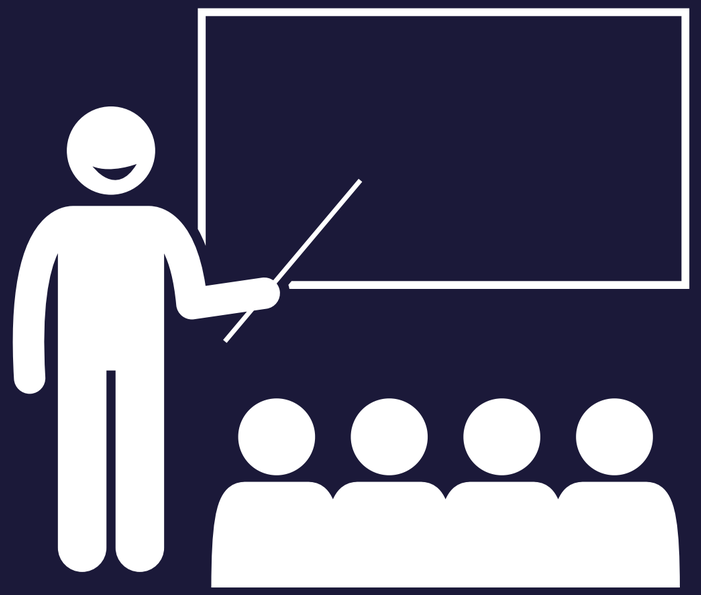
UNIVERSITY OF KANSAS

# LOOKUPS



010

|UNIVERISTY OF KANSAS



# INSTRUCTOR DEMONSTRATION

Lookups

UNIVERSITY OF KANSAS

# LOOK IT UP WITH LOOKUPS



**Q:** **Assume this table is gigantic. How would we retrieve the population of a specific planet for use in another formula?**

| Planet | Population |
|---|---|
| Zeelo | 5020 |
| Merinoa | 380 |
| Cardboard Box | 2 |
| ... | ... |
| Asteroid 9 | 95 |

# LOOK IT UP WITH LOOKUPS



**Q:** Assume this table is gigantic. How would we retrieve the population of a specific planet for use in another formula?

**A:** `=vlookup( <value>, <full table>, <column to retrieve>,<match parameter>)`

| Planet | Population |
|---|---|
| Zeelo | 5020 |
| Merinoa | 380 |
| Cardboard Box | 2 |
| ... | ... |
| Asteroid 9 | 95 |

# LOOK IT UP WITH LOOKUPS



=VLOOKUP(What you want to look up, where you want to look for it, the column number containing the value to return, return an Approximate or Exact match – indicated as 1/TRUE, or 0/FALSE).

# LOOK IT UP WITH LOOKUPS



**Q:** What will this yield?

```
=vlookup( “Asteroid 9”, Planets, 3, FALSE)
```

**Planets Table**

| Planet | Population | Species |
|---|---|---|
| Zeelo | 5020 | Zoltans |
| Merinoa | 380 | Murphies |
| Cardboard Box | 2 | Hambones |
| … | … | |
| Asteroid 9 | 95 | Asterisks |

# LOOK IT UP WITH LOOKUPS



**Q:** What will this yield?
`=vlookup( “Asteroid 9”, Planets, 3, FALSE)`

**A:** Asterisks

| Planet | Population | Species |
|---|---|---|
| Zeelo | 5020 | Zoltans |
| Merinoa | 380 | Murphies |
| Cardboard Box | 2 | Hambones |
| … | … | |
| Asteroid 9 | 95 | Asterisks |

# ACTIVITY: PRODUCT PIVOT

A small company selling electronics and electronic media has asked our class to create a table that visualizes the cost of their recent orders. Using lookups, create a pivot table that serves this purpose.

**Suggested Time: 15 minutes**

UNIVERSITY OF KANSAS

# PRODUCT PIVOT INSTRUCTIONS



- Determine the “Product Price” of each row in the “Orders” sheet by using a VLOOKUP() that references each row's “Product ID.”

  The “Product Price” of a row does not include shipping.

- Determine the “Shipping Price” of each row in the “Orders” sheet by using a VLOOKUP() that references each row's “Shipping Priority.”

- Select all of the data on the “Orders” sheet and create a new pivot table that calculates the sum of both “Product Price” and “Shipping Price” for each “Order Number” and “Product ID.”

**Suggested Time: 15 minutes**

# CHECK IN

Let's Review

UNIVERSITY OF KANSAS

# SUMMARY

- Calculating averages was covered in Lesson 1.2.5.
- Creating pivot tables was covered in Lesson 1.3.1.
- Using the VLOOKUP() function was covered in Lesson 1.4.2