

Data Cleaning in Introductory Accounting: What, Why, and How

Jennifer Cainas, DBA, CPA
Tracie Miller, Ph.D., CPA
Wendy Tietz, Ph.D., CMA, CSCA, CPA, CGMA, MCT

September 16, 2022



Webinar Features and CPE Credit



Asking Questions



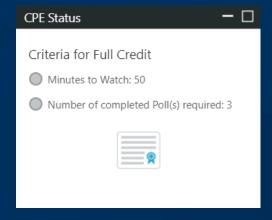
Closed Captioning



Help



CPE Credit





Moderator

Jodi L. Ryan, CAE
Director, Member Engagement
IMA





Featured Presenter

Jennifer Cainas, DBA, CPA
Associate Dean and
Professor of Instruction
University of South Florida





Jennifer M. Cainas's Biography

- Jennifer M. Cainas, DBA, CPA, is a professor of instruction and associate dean and the Muma College of Business at the University of South Florida. Dr. Cainas primarily teaches introductory financial and managerial accounting, but has also taught Intermediate Financial Accounting, Cost Accounting, and Accounting Information Systems. During her teaching career, she has taught and mentored over 25,000 business students and has won numerous teaching awards at the college, university, state, and national level.
- She is a member of the American Accounting Association, the Institute of Management Accountants, and the Florida Institute of Certified Public Accountants. She has been very active within the American Accounting Association, most recently serving on the Management Accounting Section Teaching Symposium Committee (2021-2022).
- Jennifer is a co-author of a managerial accounting textbook with Pearson Education. She is passionate about student success and engagement, as well as helping other faculty members develop effective teaching methods.

Featured Presenter

Tracie Miller, Ph.D., CPA

Lead Faculty

Franklin University





Tracie L. Miller's Biography

- Tracie L. Miller (Nobles), Ph.D., CPA, is the lead faculty for the B.S. Accounting program at Franklin University. She has extensive teaching experience including faculty appointments at Austin Community College, Texas State University, Aims Community College, and The University of Texas-Austin.
- She serves in numerous leadership roles for American Accounting Association, American Institute of CPAs, Teachers of Accounting at Two-Year Colleges, and Texas Society of CPAs. Dr. Miller received her doctoral degree from Texas A&M University. Her research interest includes financial literacy education, culturally responsive teaching, and online learning. She has public accounting experience with Deloitte Tax LLP.
- Dr. Miller has received several teaching and professional awards including the AAA / J. Michael and Mary Anne / Deloitte Foundation Cook Prize, TXCPA Outstanding Accounting Educator, and AICPA Bea Sanders Award. Dr. Miller is also an author of several accounting textbooks published by Pearson.

Featured Presenter

Wendy Tietz, Ph.D., CMA, CSCA, CPA, CGMA, MCT

Professor

Kent State University





Wendy M. Tietz's Biography

- Wendy M. Tietz, PhD, CPA, CGMA, CMA, CSCA, MCT, is an accounting professor at Kent State University where she teaches introductory financial and managerial accounting in a variety of formats as well as advanced management accounting and advanced accounting analytics and technology. She has taught more than 20,000 students to date and has received several teaching awards, including the American Accounting Association Cook Deloitte Foundation Prize for undergraduate teaching, and the IMA R. Lee Brummet Exemplary Educator Award.
- Dr. Tietz is active in the Institute of Management Accounting, the American Accounting Association, and several other professional accounting organizations. She is also the coauthor of two introductory accounting textbooks with Pearson Education. She is passionate about teaching accounting and using educational technology to enhance student engagement.

Learning Objectives

- 1. Define data cleaning (ETL).
- 2. Describe why data cleaning is important in introductory accounting.
- 3. Apply a data cleaning project to an accounting course.

Poll Question 1

What classes do you teach? (Please select all that apply)

- a. Financial accounting
- b. Managerial accounting
- c. Other accounting
- d. Not teaching currently



Poll Question 1 Results: (Placeholder)

What is ETL?

First step in data analysis involves preparing the data

- Extract obtain the data
- Transform clean and combine data from multiple sources
- Load load data into software such as Power BI or Tableau



Why is ETL important at intro level?



Shows that accounting is not just DRs and CRs



Builds data literacy



Understand that data is not "clean"



Provides a real-world, career-ready skill

KAT Concession Supply

- KAT Concession Supply is a fictitious company
- >300 records in each data set
- Different data sets for Excel, Alteryx + Power BI, Tableau + Tableau Prep, and Google Sheets



Six possible software tools









Poll Question 2

What software tools have you used in your classes? (Choose all that apply)

- a. Excel
- b. Tableau
- c. Power BI
- d. None/NA



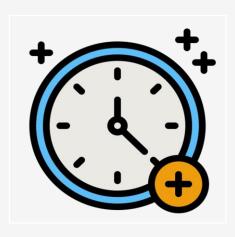
Poll Question 2 Results: (Placeholder)

Two difficulty levels



Beginner

20 - 30 minutes



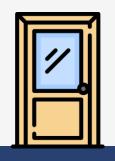
Intermediate

45 – 60 minutes

Seven possible project assignments

- 1. Excel beginner
- 2. Excel intermediate
- 3. Alteryx + Power BI beginner
- 4. Alteryx + Power Bl intermediate

- 5. Tableau + Tableau Prep beginner
- 6. Tableau + Tableau Prep intermediate
- 7. Google Sheets



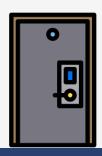












Projects

- Students download data set
- Students use real software
- Projects are preparing students for real-life with these skills of using software and basic skills such as downloading and saving files

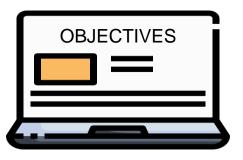
All data cleaning projects include

- Step-by-step tutorial videos
- Student guide
- Teaching notes
- Data set (editable)
- Assumes no prior software knowledge
- Can be completed in 20 60 minutes

ETL project learning objectives

After completing the project, students will be able to:

- 1. Define ETL and describe its importance
- 2. Apply data cleaning techniques
- 3. Analyze data



Why the different software tools?

- Excel is a familiar tool for most students and is accessible by most students
- Alteryx and Tableau Prep are both leading ETL tools on the market
- Power BI and Tableau are leading visualization tools
- Google Sheets is available for students who only have a Chromebook

KAT Concession Supply projects

- Start with step-by-step tutorial videos
- No prior knowledge assumed
- Each project has two parts:
 - Data cleaning
 - Dashboard creation

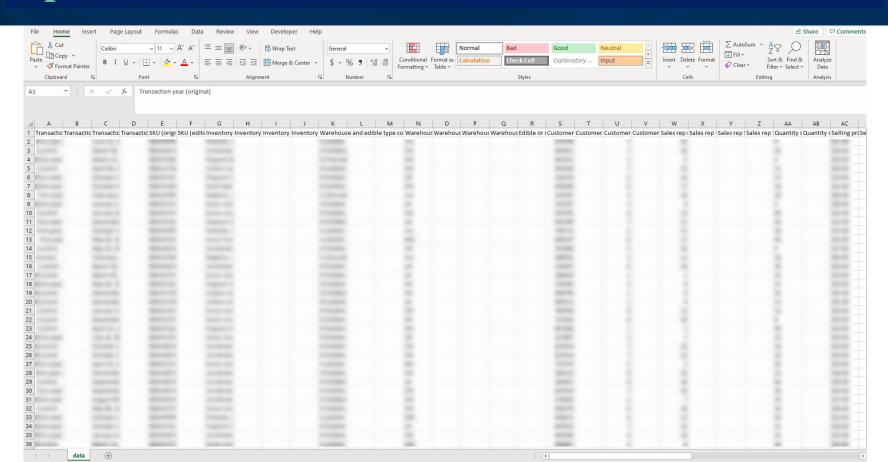
Next up, Part 1 from the Excel beginner tutorial video

No prior software knowledge assumed

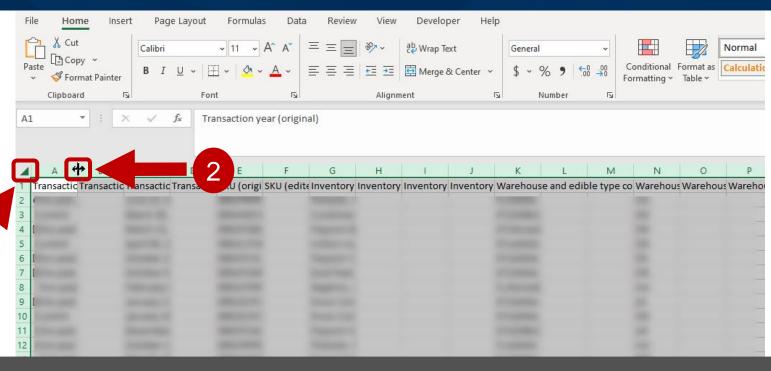
Step 01

Open Excel file and ready it for use

Open the Excel file

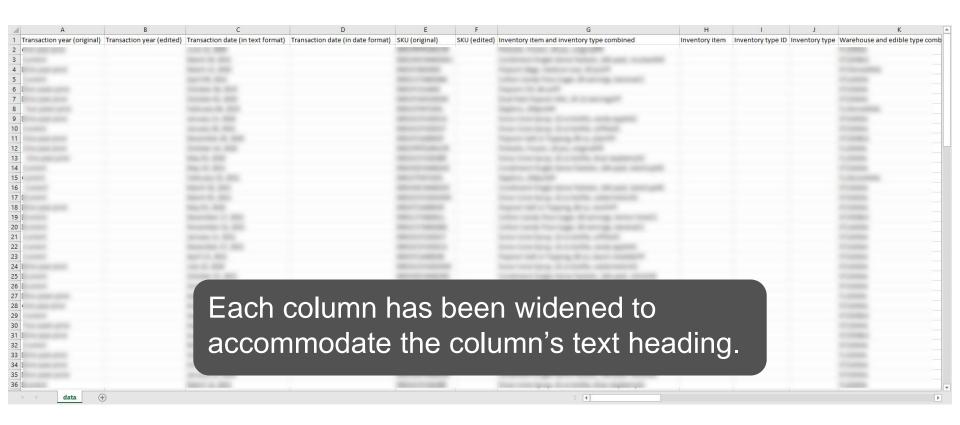


Widen all columns to fit



Click the Select All button in the top left corner. Then double-click the boundary line between the Column A and Column B headings.

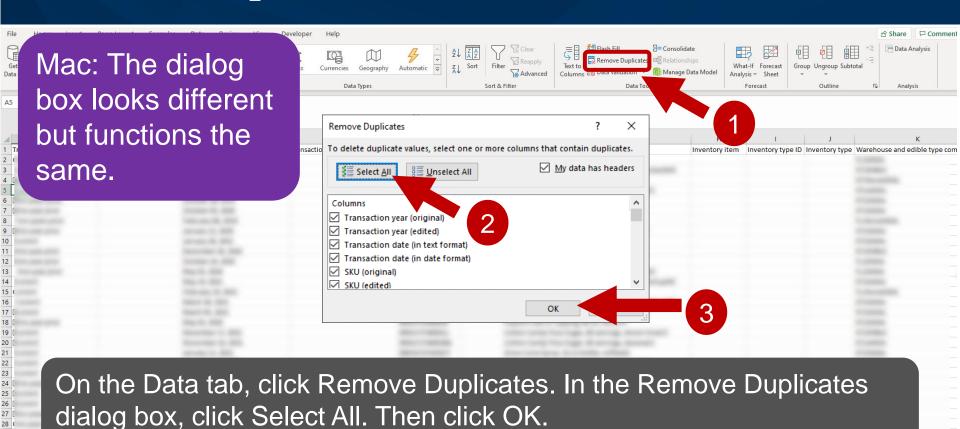
Widen all columns to fit



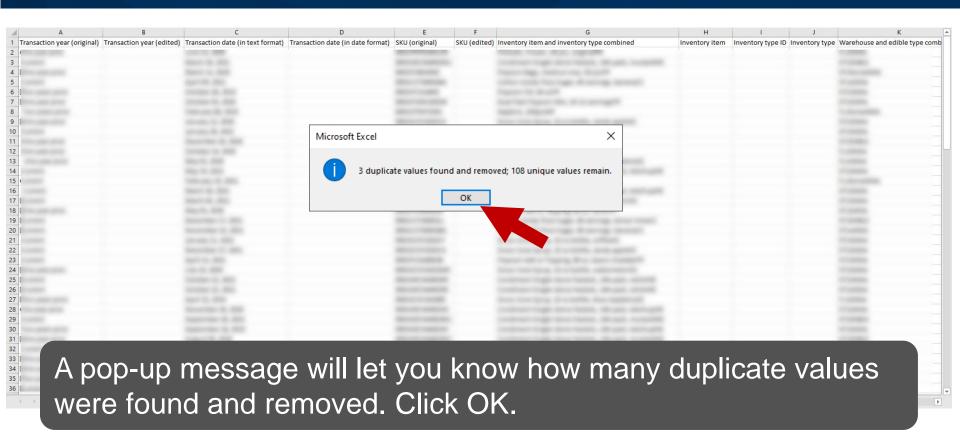
Step 02

Find and remove duplicate records

Remove duplicate records



Remove duplicate records



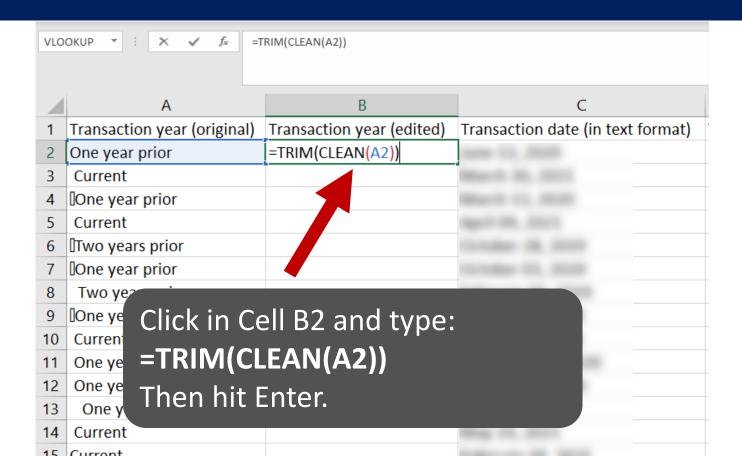
Step 03

Remove spaces and nonprinting characters from text

Remove spaces and nonprinting characters from text

	Α
1	Transaction year (original)
2	One year prior
3	Current
4	lOne year prior
5	Current
6	ITwo years prior
7	lOne year prior
8	Two years prior
9	lOne year prior
10	Current
11	One year prior
12	One year prior

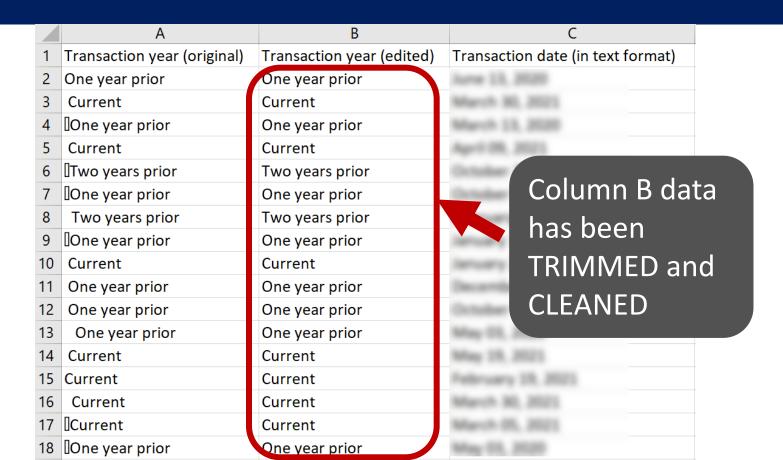
Remove spaces and nonprinting characters from text



Remove spaces and nonprinting characters from text

	Α	В	
1	Transaction year (original)	Transaction year (edited)	
2	One year prior	One year prior	
3	Current		
4	One year prior		
5	Current Double-click the	a Fill bandla	

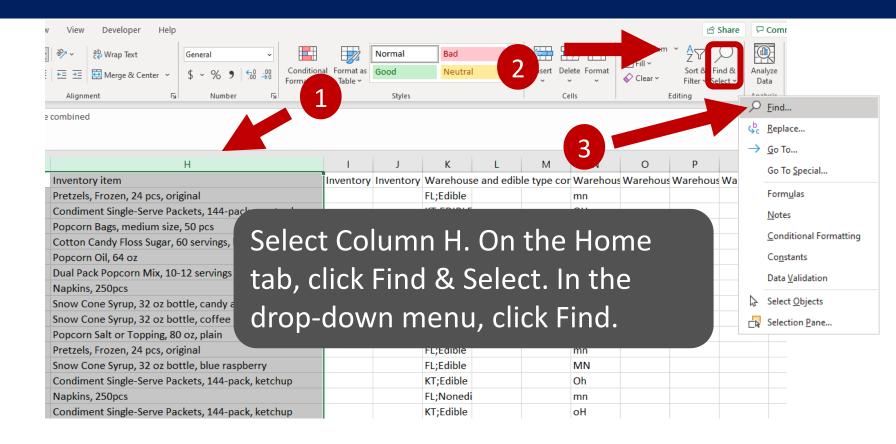
Remove spaces and nonprinting characters from text



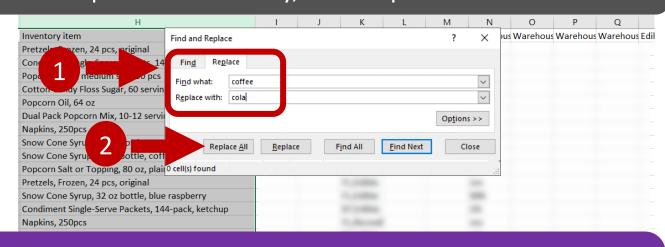
Step 04

Find and replace text

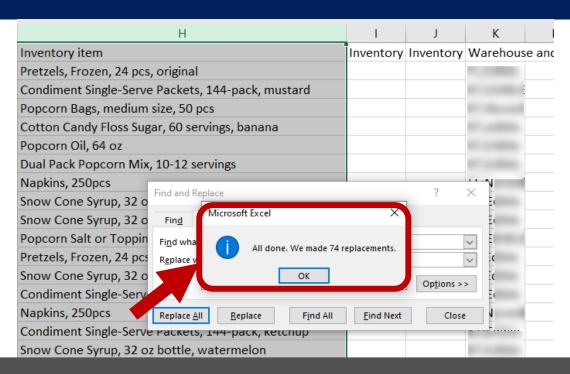
Н
Inventory item
Pretzels, Frozen, 24 pcs, original
Condiment Single-Serve Packets, 144-pack, mustard
Popcorn Bags, medium size, 50 pcs
Cotton Candy Floss Sugar, 60 servings, banana
Popcorn Oil, 64 oz
Dual Pack Popcorn Mix, 10-12 servings
Napkins, 250pcs
Snow Cone Syrup, 32 oz bottle, candy apple
Snow Cone Syrup, 32 oz bottle, coffee
Popcorn Salt or Topping, 80 oz, plain
Pretzels, Frozen, 24 pcs, original



In the Find and Replace dialog box, click the Replace tab. Then type "coffee" for Find what and type "cola" for Replace with. Finally, click Replace All.



Mac: At the top of the Find dialog box, type "coffee" for Find what. Click Replace at the bottom of the dialog box and type "cola" for Replace with. Click Replace All.



A pop-up message will let you know how many replacements were done. Click OK. Then click Close.

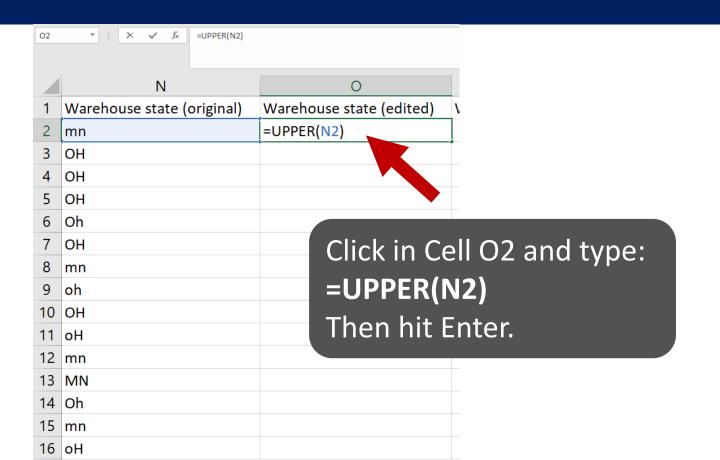
Step 05

Change a field to UPPER case

Change Warehouse state (Column N) to uppercase

	N
1	Warehouse state (original)
2	mn
3	ОН
4	ОН
5	ОН
6	Oh
7	ОН
8	mn
9	oh
10	ОН
11	оН
12	mn

Change Warehouse state (Column N) to uppercase



Change Warehouse state (Column N) to uppercase

	N	0
1	Warehouse state (original)	Warehouse state (edited)
2	mn	MN
3	ОН	ОН
4	ОН	ОН
5	ОН	ОН
6	Oh	ОН
7	ОН	ОН
8	mn	MN
9	oh	ОН
10	ОН	ОН
11	оН	ОН
12	mn	MN
13	MN	MN
14	Oh	ОН
15	mn	MN
16	оН	ОН

Double-click the Fill handle to copy the formula down the rest of the column.

Step 06

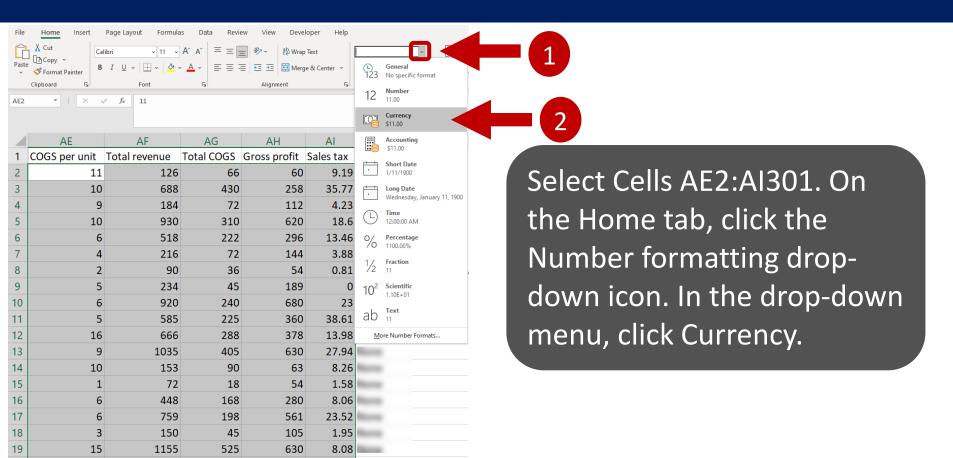
Format as currency

	AE	AF	AG	AH	Al
1	COGS per unit	Total revenue	Total COGS	Gross profit	Sales tax
2	11	126	66	60	9.19
3	10	688	430	258	35.77
4	9	184	72	112	4.23
5	10	930	310	620	18.6
6	6	518	222	296	13.46
7	4	216	72	144	3.88
8	2	90	36	54	0.81
9	5	234	45	189	0
10	6	920	240	680	23
11	5	585	225	360	38.61
12	16	666	288	378	13.98

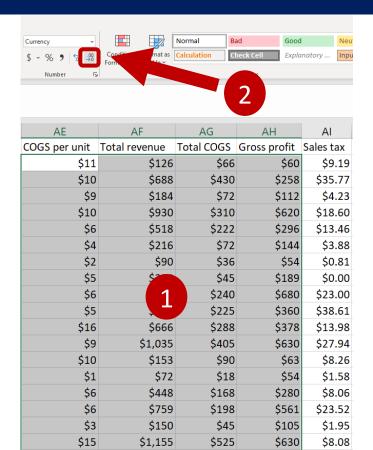
A	AE	AF	AG	AH	Al
1	COGS per unit	Total revenue	Total COGS	Gross profit	Sales tax
2	14	462	196	266	17.55
3	3	780	117	663	23.4
4	15	770	330	440	10.78
5	6	629	222	407	25.16
6	10	950	500	450	17.1
7	14	1204	602	602	40.93

Select Cells AE2:AI2. Then hit Ctrl+Shift+down arrow on your keyboard to select all the rows.

Mac: Command+Shift+down arrow



Now select Cells AE2:AH2 (we are not including Column AI this time) and hit Ctrl+Shift+down arrow to select all the rows. Then, on the Home tab, click the decrease decimal icon twice.



4	AE	AF	AG	AH	Al
1	COGS per unit	Total revenue	Total COGS	Gross profit	Sales tax
2	\$14	\$462	\$196	\$266	\$17.55
3	\$3	\$780	\$117	\$663	\$23.40
4	\$15	\$770	\$330	\$440	\$10.78
5	\$6	\$629	\$222	\$407	\$25.16
6	\$10	\$950	\$500	\$450	\$17.10
7	\$14	\$1,204	\$602	\$602	\$40.93
8	\$12	\$1,050	\$420	\$630	\$2.10
9	\$5	\$714	\$210	\$504	\$19.27
10	\$10	\$1,376	\$430	\$946	\$38.52
11	\$12	\$242	\$132	\$110	\$7.74
12	\$4	\$96	\$32	\$64	\$2.01
13	\$10	\$496	\$160	\$336	\$4.96
14	\$5	\$528	\$220	\$308	\$6.33
	***	4000	40.00	4	444.00

Now we have Columns AE through AI formatted as currency

Data set has been cleaned!



All steps are depicted in tutorial videos

No prior software knowledge assumed

Poll Question 3

What does ETL stand for?

- a. Extract, Transform, and Locate
- b. Expand, Transfer, and Load
- c. Extract, Transform, and Load
- d. Expand, Transform, and Locate



Poll Question 3 Results: (Placeholder)

7 requirements in Beginner Excel project

- 1. Widen all columns to automatically fit the contents
- 2. Find and remove duplicate records
- Remove unwanted characters such as numbers, spaces, and nonprintable characters from fields
- 4. Find and replace text with replacement text
- 5. Change the case of a text field to uppercase or lowercase
- 6. Format numbers
- Create a two-visualization interactive dashboard

Intermediate & Advanced Projects

1. Intermediate projects

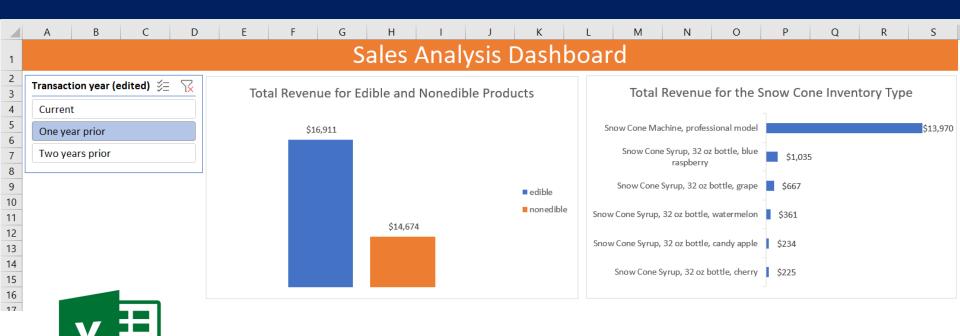
- Approximately 16 requirements
- 3-visualization dashboard

2. Advanced projects

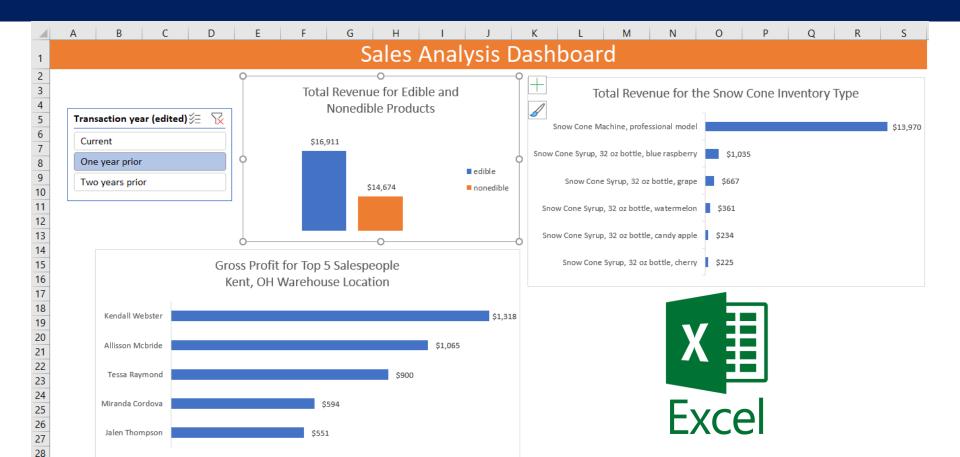
- Approximately 25 requirements
- 4-visualization dashboard
- Appropriate for honors students or for classes beyond introductory accounting

End result: Dashboards for each project

Excel beginner dashboard



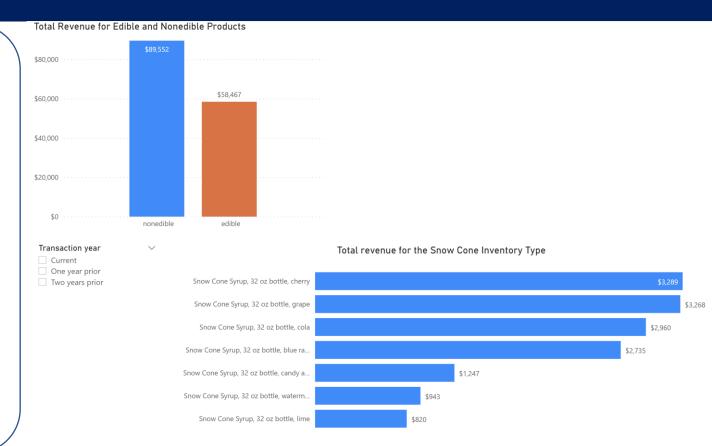
Excel intermediate dashboard



Alteryx + Power BI beginner dashboard







Alteryx + Power BI intermediate dashboard





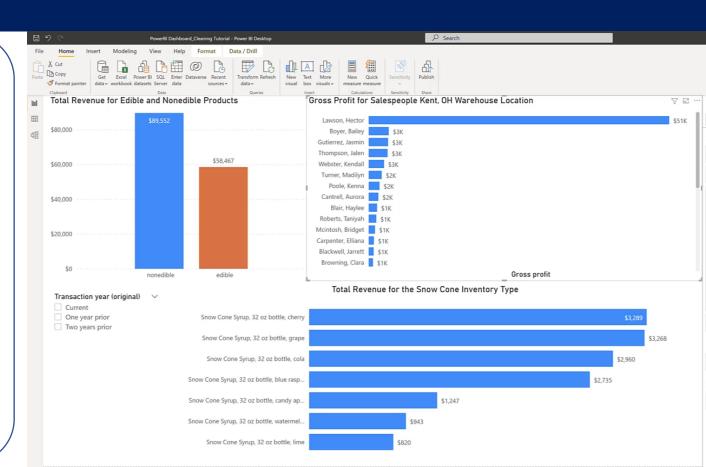
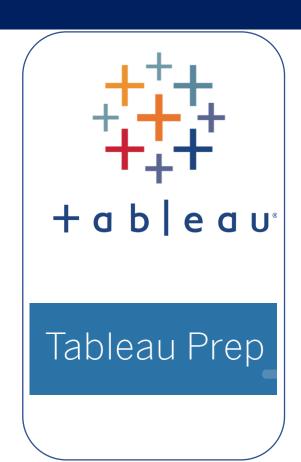


Tableau + Tableau Prep beginner dashboard



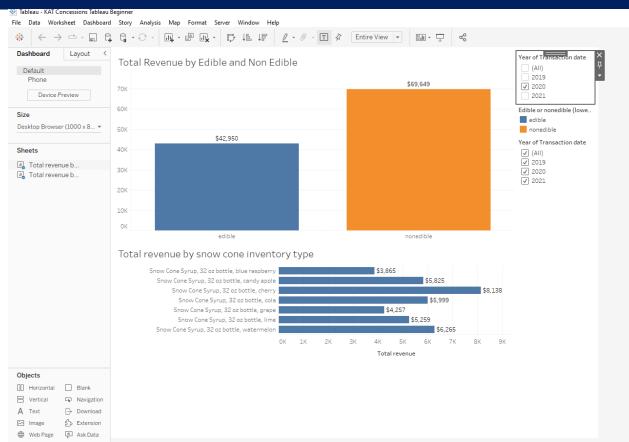
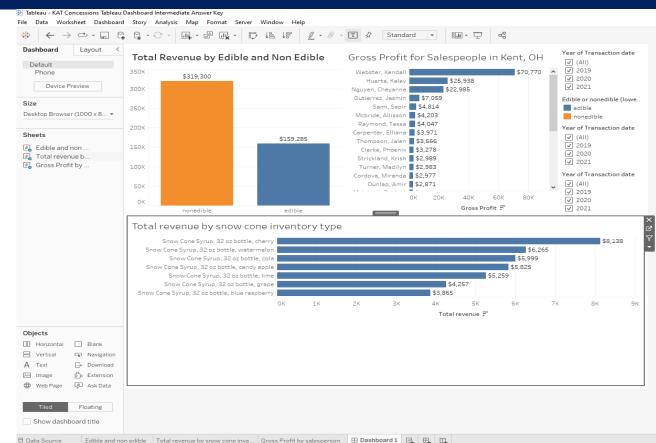


Tableau + Tableau Prep intermediate dashboard

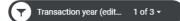


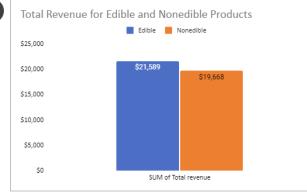
Tableau Prep

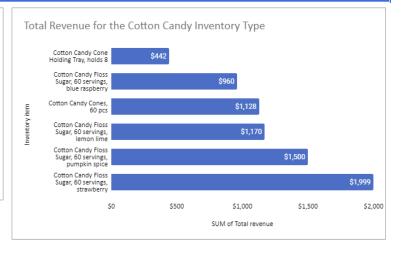


Google Sheets dashboard

Data Analysis Dashboard









Google Sheets

General teaching strategy



Step-by-step instructions



No prior knowledge needed

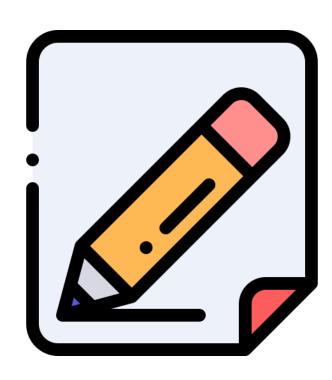
....for instructors and students

All ETL projects include

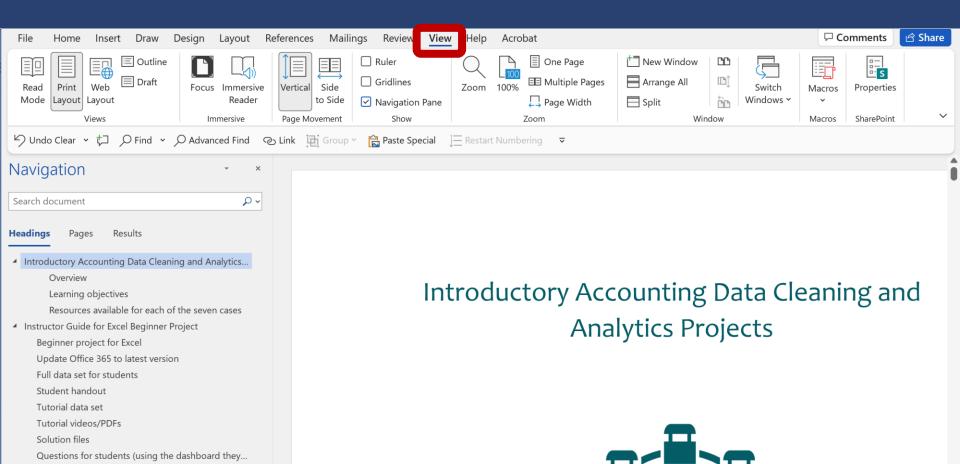
- Project student guide
- Step-by-step tutorial videos (close captioned)
- Scripted tutorial videos slides (pdf)

- Tutorial data set
- Project data set (editable)
- Discussion questions
- Solution files

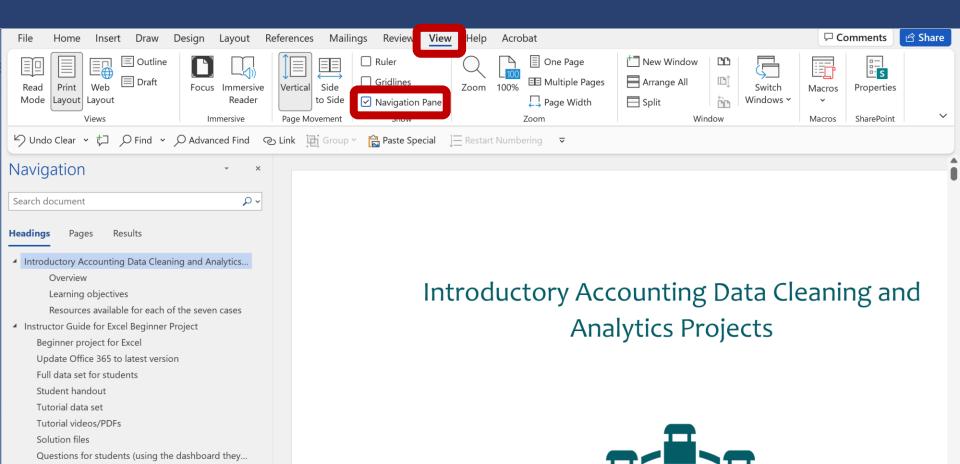
Data sets are editable



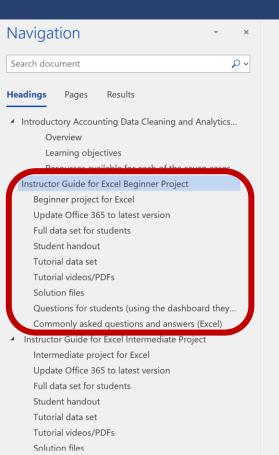
Instructor Guide: Use Navigation Pane in Word



Instructor Guide: Use Navigation Pane in Word



Instructor Guide: Use Navigation Pane in Word



Instructor Guide for Excel Beginner Project

KAT Concessions Supply is a fictitious company that supplies food, paper products, and other items to fairs, amusement parks, and other organizations. Information in the KAT Concession Supply data sets includes information about its sales over a three-year period. The data from KAT Concession Supply has errors and irregularities in it and it must be cleaned so that an interactive dashboard about its operations can be prepared.

"ETL" stands for Extract, Transform, and Load. The ETL process extracts, transforms, and loads data from multiple sources to a data warehouse (single data source) where that data can be used in decision-making. ETL is the foundation for data analytics.

To show students the "why" of ETL, an Excel interactive dashboard will be built after the data cleaning steps are completed. This beginner project includes a two-visualization dashboard.

Beginner project for Excel

The ETL and other tasks in the beginner version of the Excel data cleaning project are:

- 1. Widen all columns to automatically fit the contents
- 2. Find and remove duplicate records
- 3. Remove unwanted characters such as numbers, spaces, and nonprintable characters from fields
- 4. Find and replace text with replacement text
- 5. Change the case of a text field to uppercase or lowercase
- 6. Format numbers
- 7. Create a two-visualization interactive dashboard

Update Office 365 to latest version

To do the project with Excel, students will need to use Microsoft 365. Before starting the project, students need to update their version of Excel. For instructions on how to update the Excel application, search for "how to update Office 365" on the internet. Either Windows or Mac can be used for this project. The instructions included in the step-by-step tutorial video use Excel for Windows, but Excel for

Instructor Guide: Use Navigation Pane in Word

Navigation

Search document

Headings

▲ Introductory Acco

Overview Learning ob

Pages

nstructor Guide

Beginner proje
Update Office

Full data set fo

Student hando Tutorial data se

Tutorial videos

Solution files

Questions for s Commonly ask

Instructor Guide

Intermediate p
Update Office

Full data set for

Student hando

Solution files

Tutorial data se

Instructor Guide for Excel Beginner Project

Beginner project for Excel

Update Office 365 to latest version

Full data set for students

Student handout

Tutorial data set

Tutorial videos/PDFs

Solution files

Questions for students (using the dashboard they...

Commonly asked questions and answers (Excel)

Beginner Project

es food, paper products, and other items to ion in the KAT Concession Supply data sets ad. The data from KAT Concession Supply has an interactive dashboard about its

cess extracts, transforms, and loads data from where that data can be used in decision-

shboard will be built after the data cleaning visualization dashboard.

I data cleaning project are:

aces, and nonprintable characters from fields

ercase

rosoft 365. Before starting the project, ions on how to update the Excel application,

Tutorial videos/PDFs

search for "how to update Office 365" on the internet. Either Windows or Mac can be used for this project. The instructions included in the step-by-step tutorial video use Excel for Windows, but Excel for

General teaching strategy: Delivery

- Assigned as homework
- Students use step-by-step tutorial videos
- Beginner projects: 20 30 minutes to complete
- Full projects: 45 60 minutes to complete
- Debrief in class or online

Request instructor guide for KAT Concession Supply ETL projects at http://tiny.cc/katcs



Poll Question 4

How do you get the instructor guide for the KAT Concession Supply project?

- a. Request at http://tiny.cc/katcs
- b. Request at AccountingIsAnalytics.com
- c. Ask IMA
- d. Both a and b are correct



Poll Question 4 Results: (Placeholder)

How data analytics projects fit together

Minidashboards



Web-based; no software

Pre-built dashboards

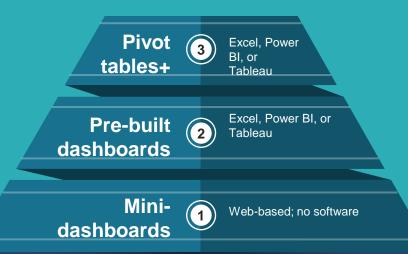


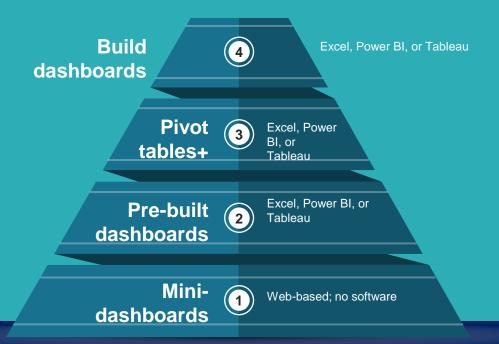
Excel, Power BI, or Tableau

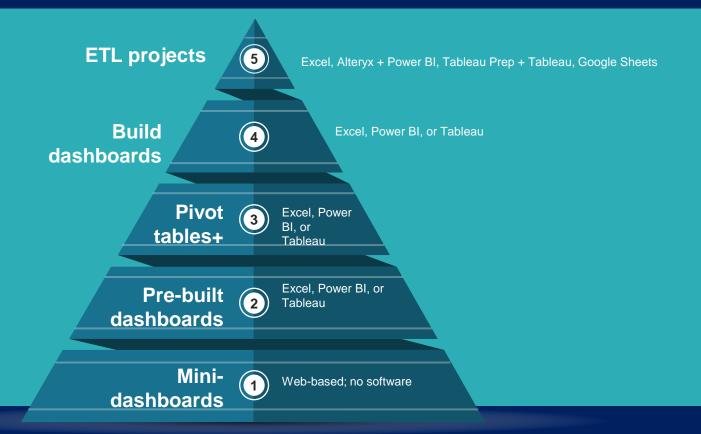
Minidashboards



Web-based; no software







Request instructor guide for KAT Concession Supply ETL projects at http://tiny.cc/katcs





For additional data analytics teaching resources, see **AccountingIsAnalytics.com**

Questions and Answers



Jennifer Cainas, DBA, CPA
Associate Dean and
Professor of Instruction
University of South Florida



Tracie Miller, Ph.D., CPA
Lead Faculty
Franklin University



Wendy Tietz, Ph.D., CMA, CSCA, CPA, CGMA, MCT Professor Kent State University



Jodi Ryan, CAE Director, Member Engagement IMA

Thank you to Our Featured Presenters!



Jennifer Cainas, DBA, CPA
Associate Dean and
Professor of Instruction
University of South Florida



Tracie Miller, Ph.D., CPA
Lead Faculty
Franklin University



Wendy Tietz, Ph.D., CMA, CSCA, CPA, CGMA, MCT Professor Kent State University

Final Reminders

- **► Complete the Evaluation poll** 2 options
 - On your screen
 - Evaluation Survey icon at the bottom of your console
- ► Access to your CPE Certificate 2 options
 - Click the "CPE" icon at the bottom of your console or
 - Click the link in your post-event e-mail
- ▶ Please print a copy of the CPE certificate for your records.
- ► Your CPE credit will be automatically recorded in your transcript.

Thank you!



The Association of Accountants and Financial Professionals in Business