Business Information
& Analytics
UNIVERSITY OF DENVER

DASI: Defining Class Expectations to Achieve Increased Student Capabilities

Adapting to GAI: Adapting Assignments

Kellie Keeling
DSI, Phoenix, AZ
November 24, 2024

bit.ly/KKPresentationFiles

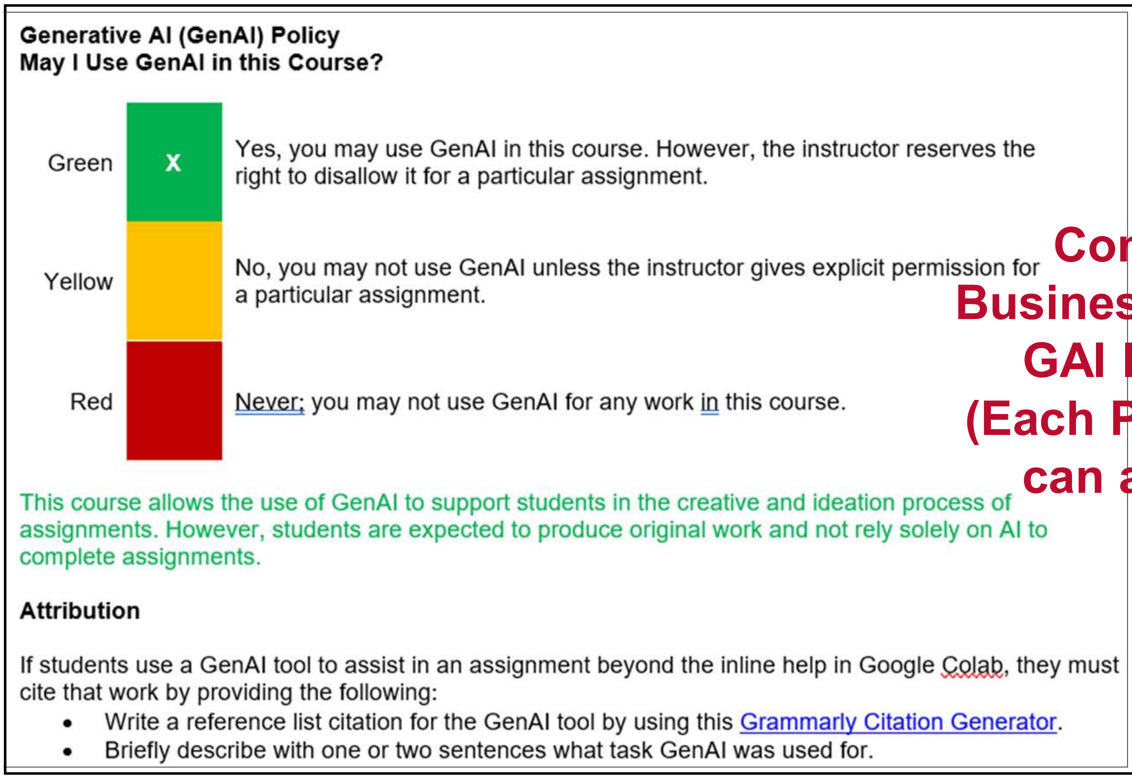
Motivation and Overview

- Easy access of GAI tools
  - Less focus on syntax – assume can code faster in R and Python
- Graduate **Predictive Modeling with R** course
  - Use R and R Studio
  - Added more questions around presenting the results from the analysis
- Undergraduate **Python Programming** class
  - Used Google Colab
  - Completely embraced the use of Gemini embedded in the interface.
  - Had them present their projects to me to help me find out if they knew what their code was doing.

bit.ly/KKPresentationFiles

Business Information
& Analytics
UNIVERSITY OF DENVER

Context Business School GAI Policy (Each Professor can adapt)



Generative AI (GenAI) Policy
May I Use GenAI in this Course?

Green	🟩 X	Yes, you may use GenAI in this course. However, the instructor reserves the right to disallow it for a particular assignment.
Yellow	🟨	No, you may not use GenAI unless the instructor gives explicit permission for a particular assignment.
Red	🟥	Never; you may not use GenAI for any work in this course.

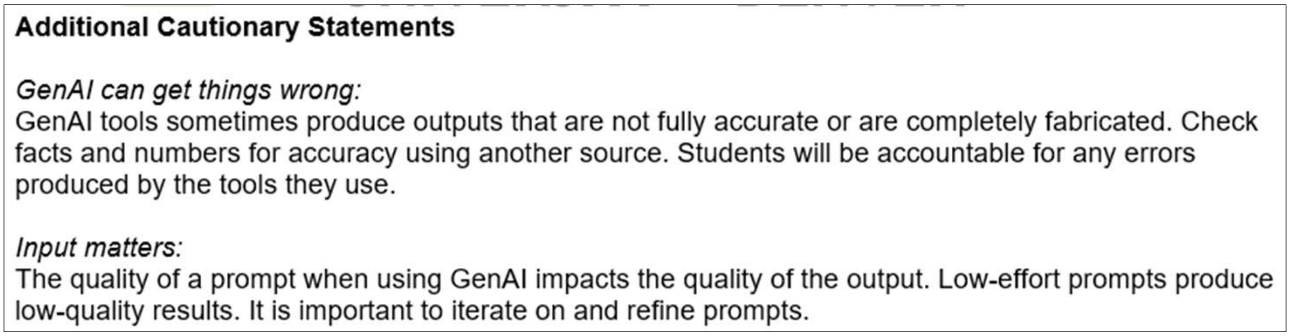
This course allows the use of GenAI to support students in the creative and ideation process of assignments. However, students are expected to produce original work and not rely solely on AI to complete assignments.

Attribution

If students use a GenAI tool to assist in an assignment beyond the inline help in Google Colab, they must cite that work by providing the following:

- Write a reference list citation for the GenAI tool by using this Grammarly Citation Generator.
- Briefly describe with one or two sentences what task GenAI was used for.

Context – Business School GAI Policy (Each Professor can adapt)



Additional Cautionary Statements

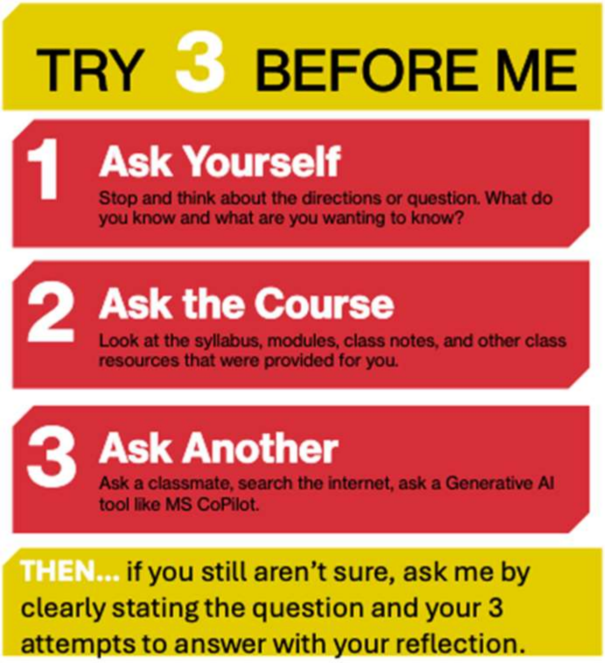
GenAI can get things wrong:
GenAI tools sometimes produce outputs that are not fully accurate or are completely fabricated. Check facts and numbers for accuracy using another source. Students will be accountable for any errors produced by the tools they use.

Input matters:
The quality of a prompt when using GenAI impacts the quality of the output. Low-effort prompts produce low-quality results. It is important to iterate on and refine prompts.

The Dean has recently said we should just allow students to use GAI……

Business Information
& Analytics
UNIVERSITY OF DENVER

Context



TRY 3 BEFORE ME

1 Ask Yourself
Stop and think about the directions or question. What do you know and what are you wanting to know?

2 Ask the Course
Look at the syllabus, modules, class notes, and other class resources that were provided for you.

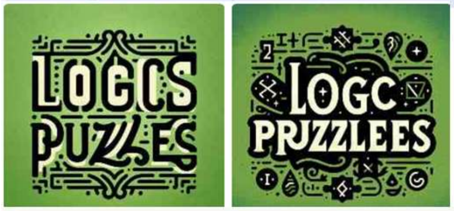
3 Ask Another
Ask a classmate, search the internet, ask a Generative AI tool like MS CoPilot.

THEN... if you still aren't sure, ask me by clearly stating the question and your 3 attempts to answer with your reflection.

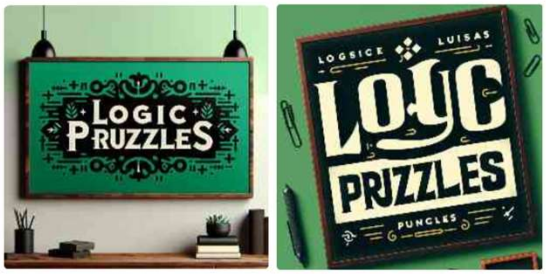
Business Information & Analytics
UNIVERSITY OF DENVER

Brief Diversion:
Great at Coding-Bad at Spelling

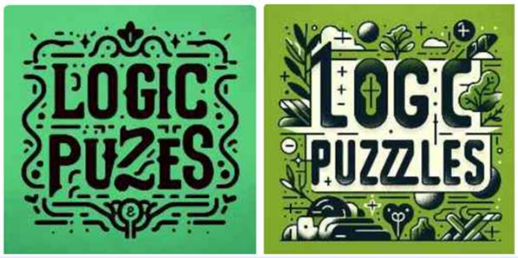
Can you make a text banner that says logic puzzles with a fancy font that is black colored with a green background?



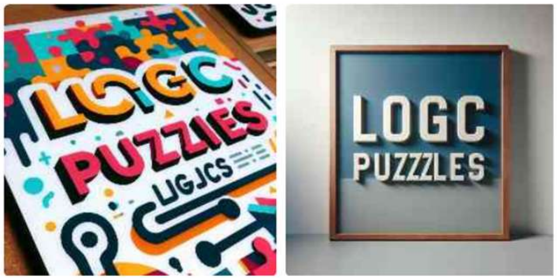
Can you fix the spelling. It should be "LOGIC PUZZLES"



Try again to fix the second work to be "PUZZLES"



Can you create a sign with the words "LOGIC PUZZLES" on it?



can you make the letters LOGIC PUZZLES with a fancy background

LOGIC
PUZZLES
LOGIGG
PUZZIES

12th Try!

can you put these 2 words on a background of snowflakes

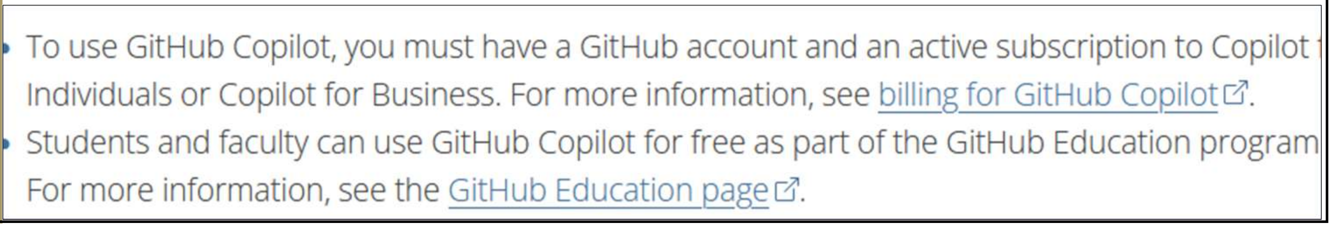
LOGC
PUZZLES
LOGIC
PUZZLES

Brief Diversion:
Great at Coding
Bad at Spelling

Business Information
& Analytics
UNIVERSITY OF DENVER

Advanced Predictive Modeling with R

- MSBA Students in their final Quarter
  - Adding more presenting results
  - Used R and R Studio
    - There is a GitHub – I couldn't get an account ☹
      - Has Anyone got this to work?
    - https://docs.posit.co/ide/user/ide/guide/tools/copilot.html



- To use GitHub Copilot, you must have a GitHub account and an active subscription to Copilot
Individuals or Copilot for Business. For more information, see billing for GitHub Copilot.
- Students and faculty can use GitHub Copilot for free as part of the GitHub Education program
For more information, see the GitHub Education page.

Course Topics: (Meets 2 2hr10m Classes per Week)

0	Complete DataCamp Courses: (Intro to Tidyverse/Intermediate R)
1	R/R Studio Intro/Tidyverse / Control Logic / Statistical Learning
2	Simple/Multiple Linear Regression
3	Regression Assumptions & Transformations / Logistic Regression
4	Discriminant Analysis / Cross Validation/Bootstrap
5	Ridge Regression and Lasso / Smoothers: Polynomials and Splines
6	Gen Additive Models / CART Trees
7	Bagging/Random Forest/Boosting
8	Support Vector Machines / Neural Networks
9	Deep Learning / Present Projects

Advanced Predictive Modeling with R

- Assignment Examples
  - Shared document for interpretation summaries
  - Expectation: more variety and allowed to see other examples
  - Example 1: Regression pruned/interaction
  - Example 2: Logistic and Discriminant Analysis
  - Final Project Presentation to Peers (present twice)

Business Information
& Analytics
UNIVERSITY OF DENVER

Advanced Predictive Modeling with R

- Example 1: Regression pruned/ interaction

- (Describe process)

Reg_adv_full: sales = 2.939 + 0.046(TV) + 0.189(radio) + -0.001(newspaper)

Reg_adv_pruned: sales = 2.921 + 0.046(TV) + 0.188(radio)
Reg_adv_pruned r^2: 0.897
Reg_adv_pruned MAE: 1.254

Reg_adv_interact: sales = 6.75 + 0.019(TV) + 0.029(radio) + 0.001(TV*radio)
Reg_adv_interact r^2: 0.968
Reg_adv_interact MAE: 0.660

The three models above were evaluated on their accuracy in predicting sales for every dollar spent in each mode of media. The newspaper predictor was removed due to being an insignificant predictor, resulting in the creation of the pruned model. A third model was created that multiplied the TV and radio predictors to test how forcing them to interact would affect the model's accuracy. The best performing model proved to be the interact model as it has the best r^2 and the lowest MAE.

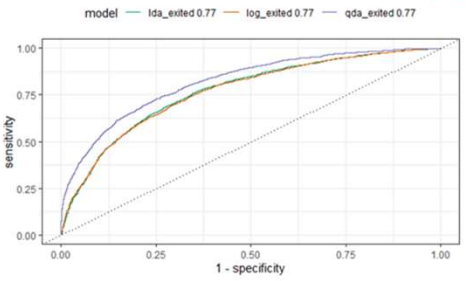
Advanced Predictive Modeling with R

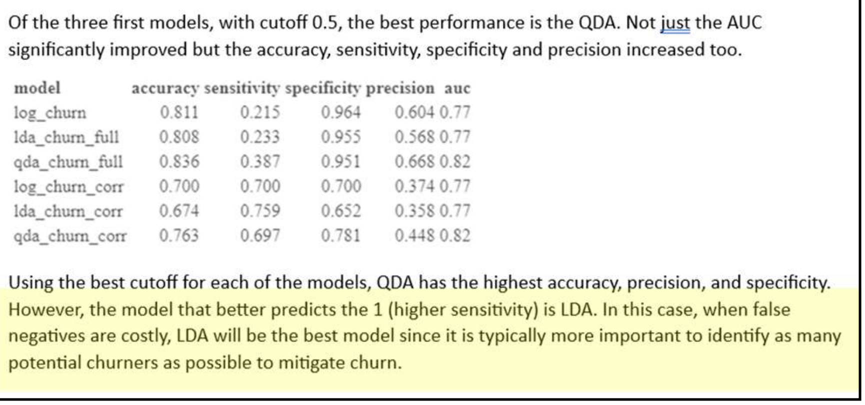
- Example 2: Logistic and Discriminant Analysis
- Include in your discussion the 5 metrics and the ROC Curve as well as the changed cutoff models. Make a case for prioritizing sensitivity or precision.

Business Information & Analytics
UNIVERSITY OF DENVER

Advanced Predictive Modeling with R

- Example 2: Logistic and Discriminant Analysis



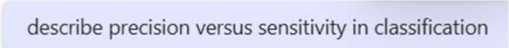


Of the three first models, with cutoff 0.5, the best performance is the QDA. Not just the AUC significantly improved but the accuracy, sensitivity, specificity and precision increased too.

model	accuracy	sensitivity	specificity	precision	auc
log_churn	0.811	0.215	0.964	0.604	0.77
lda_churn_full	0.808	0.233	0.955	0.568	0.77
qda_churn_full	0.836	0.387	0.951	0.668	0.82
log_churn_corr	0.700	0.700	0.700	0.374	0.77
lda_churn_corr	0.674	0.759	0.652	0.358	0.77
qda_churn_corr	0.763	0.697	0.781	0.448	0.82

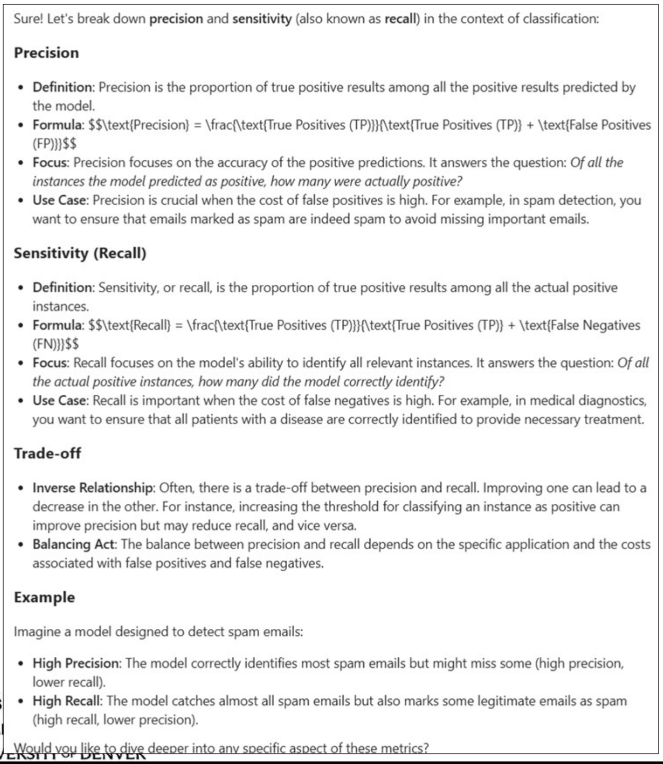
Using the best cutoff for each of the models, QDA has the highest accuracy, precision, and specificity. However, the model that better predicts the 1 (higher sensitivity) is LDA. In this case, when false negatives are costly, LDA will be the best model since it is typically more important to identify as many potential churners as possible to mitigate churn.

So did anyone think, What is precision again?



describe precision versus sensitivity in classification

But I digress…..



Sure! Let's break down **precision** and **sensitivity** (also known as **recall**) in the context of classification:

Precision

- **Definition**: Precision is the proportion of true positive results among all the positive results predicted by the model.
- **Formula**: $$\text{Precision} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Positives (FP)}}$$
- **Focus**: Precision focuses on the accuracy of the positive predictions. It answers the question: *Of all the instances the model predicted as positive, how many were actually positive?*
- **Use Case**: Precision is crucial when the cost of false positives is high. For example, in spam detection, you want to ensure that emails marked as spam are indeed spam to avoid missing important emails.

Sensitivity (Recall)

- **Definition**: Sensitivity, or recall, is the proportion of true positive results among all the actual positive instances.
- **Formula**: $$\text{Recall} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Negatives (FN)}}$$
- **Focus**: Recall focuses on the model's ability to identify all relevant instances. It answers the question: *Of all the actual positive instances, how many did the model correctly identify?*
- **Use Case**: Recall is important when the cost of false negatives is high. For example, in medical diagnostics, you want to ensure that all patients with a disease are correctly identified to provide necessary treatment.

Trade-off

- **Inverse Relationship**: Often, there is a trade-off between precision and recall. Improving one can lead to a decrease in the other. For instance, increasing the threshold for classifying an instance as positive can improve precision but may reduce recall, and vice versa.
- **Balancing Act**: The balance between precision and recall depends on the specific application and the costs associated with false positives and false negatives.

Example

Imagine a model designed to detect spam emails:

- **High Precision**: The model correctly identifies most spam emails but might miss some (high precision, lower recall).
- **High Recall**: The model catches almost all spam emails but also marks some legitimate emails as spam (high recall, lower precision).

Would you like to dive deeper into any specific aspect of these metrics?

Bus
& A
UNIVERSITY OF DENVER

Advanced Predictive Modeling with R

- Final Project Present to peers in small groups
  - Present twice
  - Second time typically faster, focuses in on results quicker
  - Should they present to me 3rd time recording to Zoom?

Python Programming (and Excel VBA)

- BSBA Students (Soph,Junior,Senior)
  - First course in major/minor
  - Know some will use GAI – Level Playing Field
    - Google Colab
    - https://colab.research.google.com

Course Topics: (Meets 2 1hr50m Classes per Week)

1	PY1: Intro to Python/Google Colab, Python Basics, User Input and Output
2	PY2: Basic Control Logic Using Python/Pseudocoding
3	PY3: Data Structures and Pandas Reading in Data
4	PY4:Functions, Cleaning, and Creating Summaries
5	PY5:Creating Summaries and Merging/Grouping Data
6	PY6:DataFrame Audit and Creating Graphs
7	PY7.Basic Statistics (t tests, ANOVA, Regression)
8	VBA1:Good Spreadsheets Modeling/Excel Basics & Built in Functions
9	VBA2: Structured Programming/Control Logic/Using Buttons/Message and Input Boxes/Worksheet Ranges/Charts
10	VBA3: Coding Wrap up and Excel/Python Communication
Finals	Project create Python program with menu to analyze data/Excel Workbook to create descriptive summaries and bar chart - Presentations to Me

Google Colab: colab.research.google.com

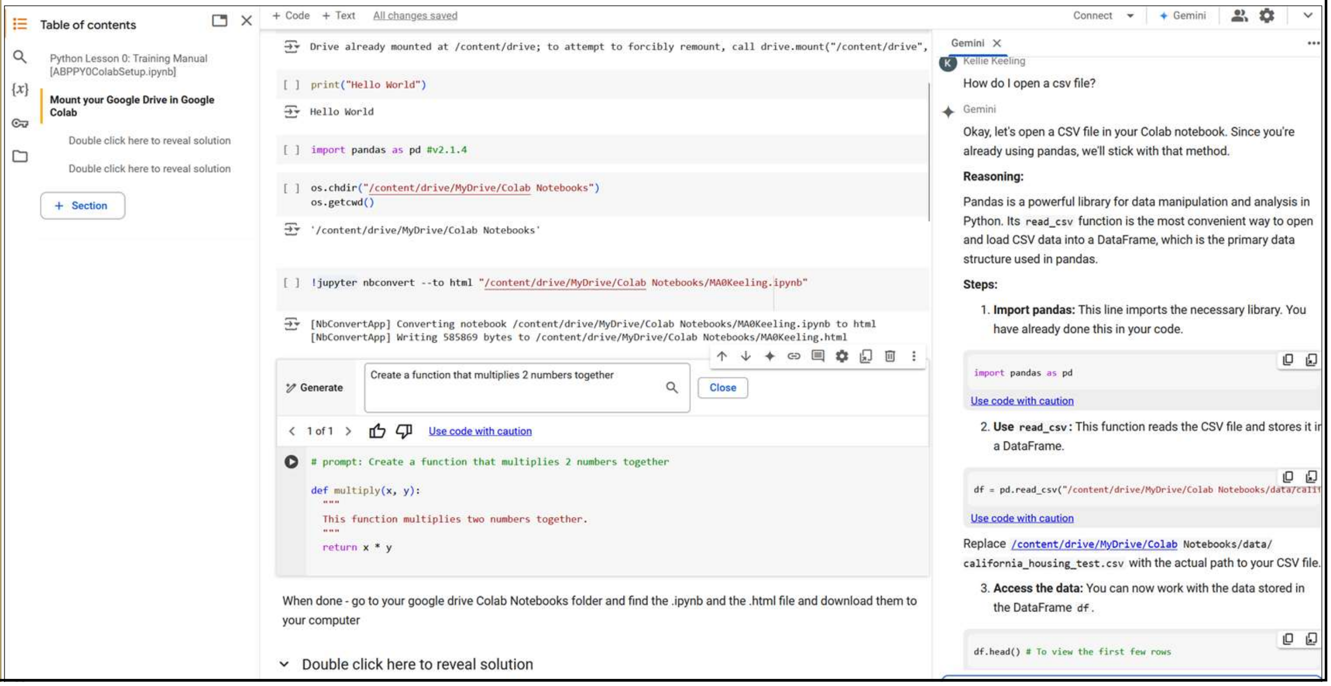


Table of contents
Python Lesson 0: Training Manual [ABPPY0ColabSetup.ipynb]
Mount your Google Drive in Google Colab
Double click here to reveal solution
Double click here to reveal solution
+ Section
+ Code
+ Text
All changes saved
Connect
Gemini
Generate
Close
1 of 1
Use code with caution
Gemini

```
Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive",
```

```
print("Hello World")
```

```
Hello World
```

```
import pandas as pd #v2.1.4
```

```
os.chdir("/content/drive/MyDrive/Colab Notebooks")
os.getcwd()
```

```
'/content/drive/MyDrive/Colab Notebooks'
```

```
!jupyter nbconvert --to html "/content/drive/MyDrive/Colab Notebooks/MA0Keeling.ipynb"
```

```
[NbConvertApp] Converting notebook /content/drive/MyDrive/Colab Notebooks/MA0Keeling.ipynb to html
[NbConvertApp] Writing 585869 bytes to /content/drive/MyDrive/Colab Notebooks/MA0Keeling.html
```

Create a function that multiplies 2 numbers together

```
# prompt: Create a function that multiplies 2 numbers together

def multiply(x, y):
  """
  This function multiplies two numbers together.
  """
  return x * y
```

When done - go to your google drive Colab Notebooks folder and find the .ipynb and the .html file and download them to your computer

Double click here to reveal solution

Kellie Keeling

How do I open a csv file?

Gemini

Okay, let's open a CSV file in your Colab notebook. Since you're already using pandas, we'll stick with that method.

Reasoning:

Pandas is a powerful library for data manipulation and analysis in Python. Its `read_csv` function is the most convenient way to open and load CSV data into a DataFrame, which is the primary data structure used in pandas.

Steps:

1. **Import pandas:** This line imports the necessary library. You have already done this in your code.

```
import pandas as pd
```

Use code with caution

2. **Use `read_csv`:** This function reads the CSV file and stores it i a DataFrame.

```
df = pd.read_csv("/content/drive/MyDrive/Colab Notebooks/data/cali
```

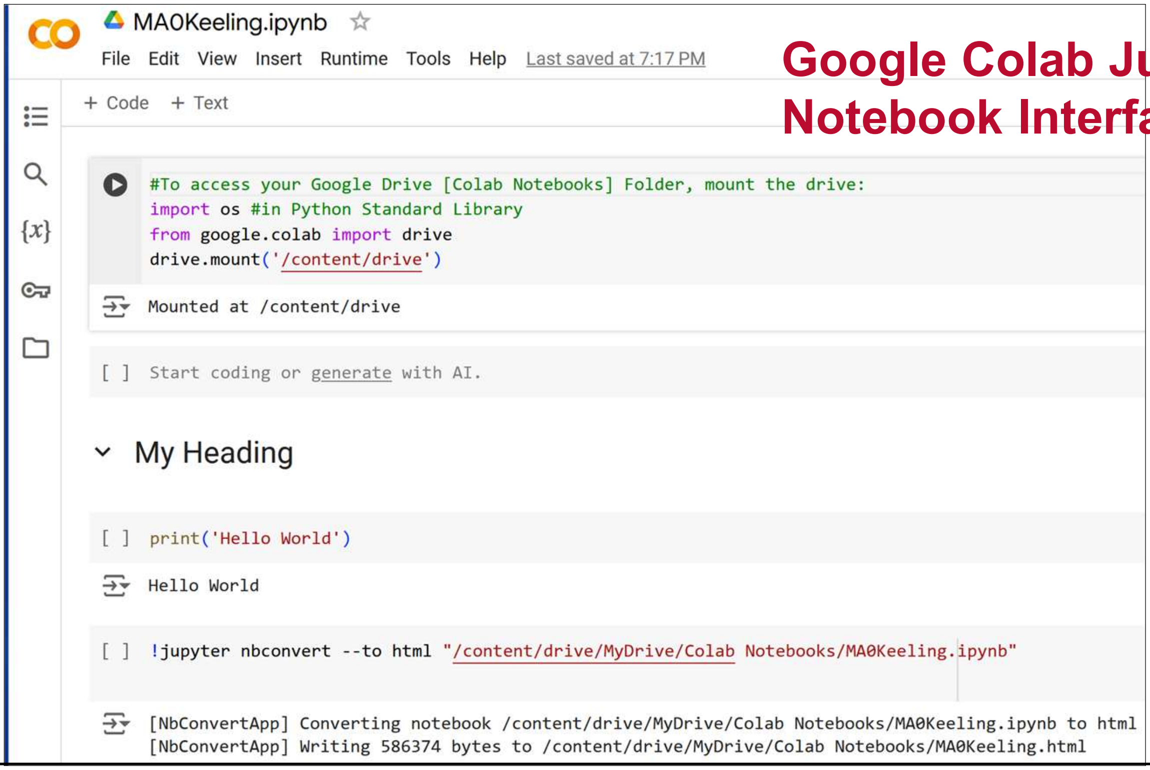
Use code with caution

Replace `/content/drive/MyDrive/Colab Notebooks/data/california_housing_test.csv` with the actual path to your CSV file.

3. **Access the data:** You can now work with the data stored in the DataFrame `df`.

```
df.head() # To view the first few rows
```

Google Colab Jupyter Notebook Interface


MA0Keeling.ipynb
File Edit View Insert Runtime Tools Help Last saved at 7:17 PM
+ Code + Text

```
#To access your Google Drive [Colab Notebooks] Folder, mount the drive:
import os #in Python Standard Library
from google.colab import drive
drive.mount('/content/drive')
```

```
Mounted at /content/drive
```

[ ] Start coding or generate with AI.

My Heading

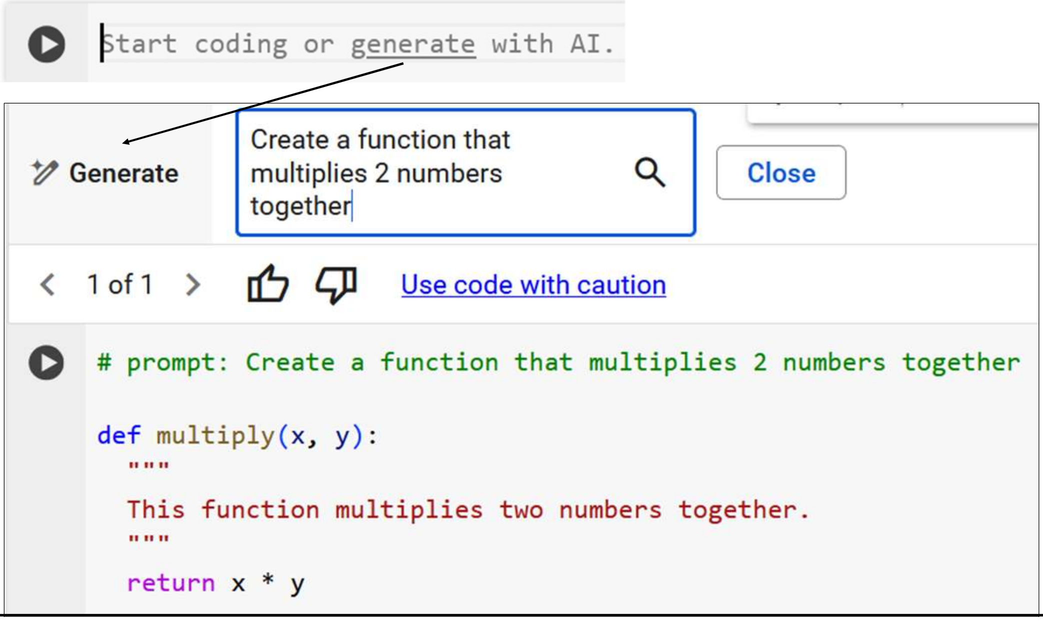
```
[ ] print('Hello World')
```

```
Hello World
```

```
[ ] !jupyter nbconvert --to html "/content/drive/MyDrive/Colab Notebooks/MA0Keeling.ipynb"
```

```
[NbConvertApp] Converting notebook /content/drive/MyDrive/Colab Notebooks/MA0Keeling.ipynb to html
[NbConvertApp] Writing 586374 bytes to /content/drive/MyDrive/Colab Notebooks/MA0Keeling.html
```

Google Colab – Generate Code

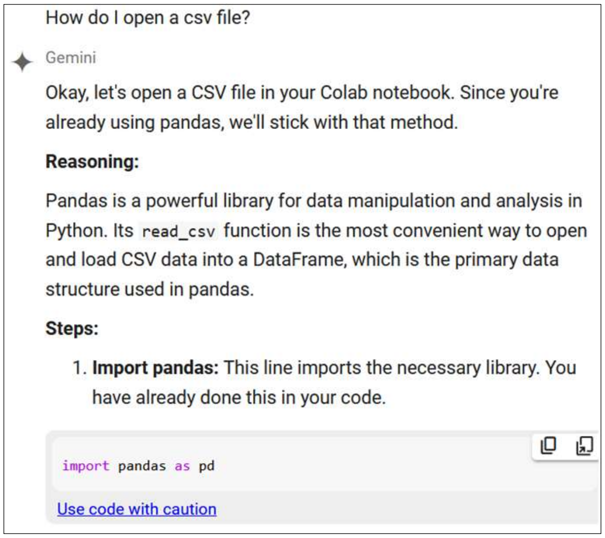


Create a function that multiplies 2 numbers together

```
# prompt: Create a function that multiplies 2 numbers together

def multiply(x, y):
  """
  This function multiplies two numbers together.
  """
  return x * y
```

Google Colab Gemini Built in



How do I open a csv file?

Gemini

Okay, let's open a CSV file in your Colab notebook. Since you're already using pandas, we'll stick with that method.

Reasoning:

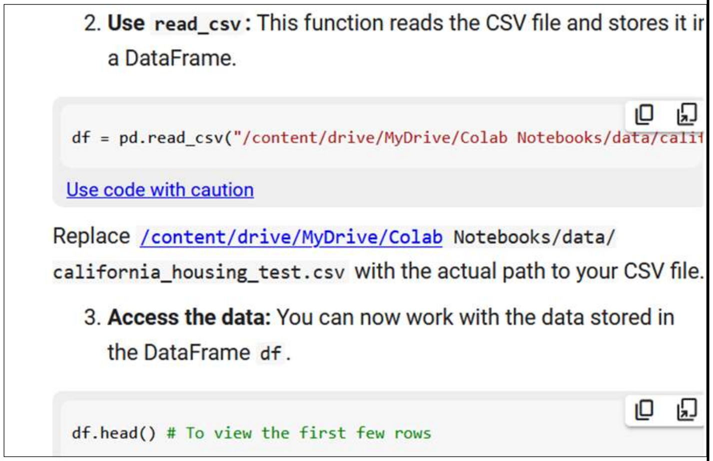
Pandas is a powerful library for data manipulation and analysis in Python. Its `read_csv` function is the most convenient way to open and load CSV data into a DataFrame, which is the primary data structure used in pandas.

Steps:

1. **Import pandas:** This line imports the necessary library. You have already done this in your code.

```
import pandas as pd
```

Use code with caution



2. **Use `read_csv`:** This function reads the CSV file and stores it i a DataFrame.

```
df = pd.read_csv("/content/drive/MyDrive/Colab Notebooks/data/calif
```

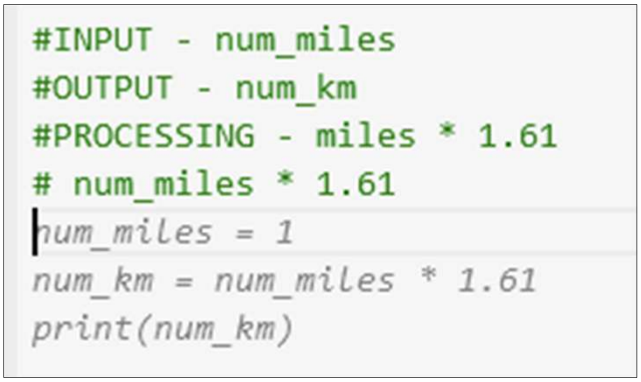
Use code with caution

Replace `/content/drive/MyDrive/Colab Notebooks/data/california_housing_test.csv` with the actual path to your CSV file.

3. **Access the data:** You can now work with the data stored in the DataFrame `df`.

```
df.head() # To view the first few rows
```

Google Colab/Gemini Examples



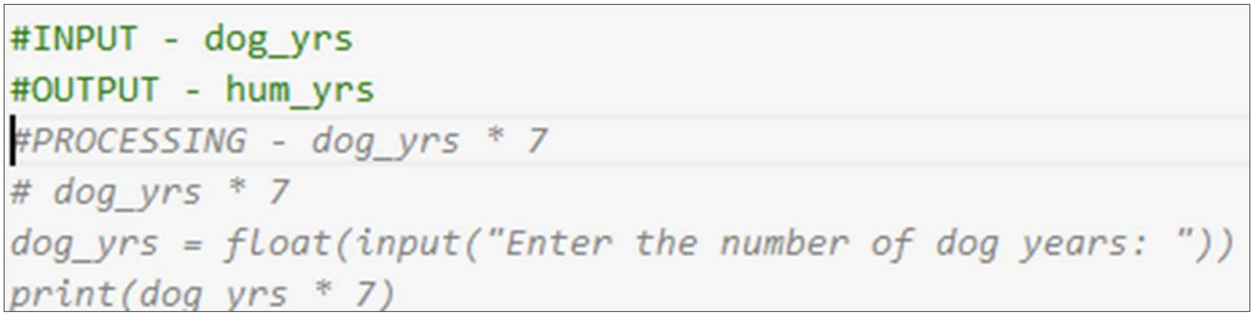
```
#INPUT - num_miles
#OUTPUT - num_km
#PROCESSING - miles * 1.61
# num_miles * 1.61
num_miles = 1
num_km = num_miles * 1.61
print(num_km)
```

I changed its suggested code to accept num_miles from user

```
num_miles = int(input(”Enter the Number of miles: ”))
```

Google Colab/Gemini Examples

Next example, I put in INPUT and OUTPUT
it scarily created the rest learning from my previous example
added the #PROCESSING
changed to using an input and converted it to float!!!



```
#INPUT - dog_yrs
#OUTPUT - hum_yrs
#PROCESSING - dog_yrs * 7
# dog_yrs * 7
dog_yrs = float(input("Enter the number of dog years: "))
print(dog yrs * 7)
```

Python Programming Class

Sample Problem:

1.Create Pseudocode WITHOUT Generative AI

Click Settings Gear on Top Right

Choose AI Assistance, **Check to Hide Generative AI Features**

Create Pseudocode to solve this Prompt with INPUTS, OUTPUTS, PROCESSING, CHECKS (create 2 checks to try out your code by finding the correct answer you should get to 2 possible entered numbers)

PROMPT: **Calculate a running sum.** Ask a user to enter a number. Keep asking for numbers until a negative number is encountered. When that happens, print the running sum of previous numbers.

2.Create Solution WITHOUT Generative AI

Using the Pseudocode you created and class notes/Python syntax file, solve the question in the prompt.

Python Programming Class

Sample Problem:

3.Create Solution WITH Generative AI

Click Settings Gear on Top Right

Choose AI Assistence, UnCheck Hide Generative AI Features

Enter the Prompt into Gemini and copy it's solution below. If you provide any additional prompts to refine the answer, include those and the revised answer as well.

4.Revise your solution from Part 2.

Copy your answer to the problem without GAI from Part 2 below.

Revise your answer based on what you learned from the Gemini Answer.

Reflect on how you changed what you did originally

Comparing Some vs None - Help

Help No Programming Exp	**Help Some Programming Exp**
Error Explanation	Time Efficiency
Advanced Code Output	Language Flexibility
Overall Helpfulness	Focus on Understanding
Learning Aid	Limited Use in Exams
Effective Use	Efficiency and Guidance
Tool Utilized	

Comparing Some vs None - Hinder

Hinder No Programming Exp	**Hinder Some Programming Exp**
Understanding Issues	Over-Reliance
Incorrect Outputs	Premature Assistance
Over-Reliance	Learning Impact
Learning Curve	Syntax Memorization
Course-Specific Impact	Future Concerns

Business Information
& Analytics
UNIVERSITY OF DENVER

Comparing Some vs None – Future Use

Future No Programming Exp	**Future Programming Exp**
Professional Applications	Documentation and Simple Tasks
Perspective and Problem-Solving	Workforce Experience
Efficiency in Basic Tasks	Decision-Making
Limitations and Ethical Considerations	Medical Applications
Programming Assistance	Educational Benefits
Selective Use	

Business Information
& Analytics
UNIVERSITY OF DENVER

Take Aways

- Encourage to turn off GAI at times – (write pseudocode first, demonstrate turning off more)
- 20 students presented projects to me this week for 12-15m (scheduled 10m). Estimated ¼ had issues explaining specific parts of code.
  - (I did pick on some things we didn't cover in class (VBA: Excel Dictionaries) or interesting features (Python: Offering list of columns for X variable that didn't include Y column chosen)
  - Wondering can this can be scaled? Or done more often?
- How else can I have them practice programming logic?

KKeeling@du.edu

bit.ly/KKPresentationFiles