National College of Ireland

Project Submission Sheet

Student Name: ……AMALACHUKWU ADAEZE ATUSIUBA……

Student ID: ……x23293012……

Programme: …. MSC AI…… **Year:** ……2024/2025…

Module: ……Programming for Artificial Intelligence (MSCAI1)……

Lecturer: ……Shreyas Setlur Arun……

Submission Due Date: ……10/12/2024……

Project Title: ……

Word Count: ……

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the references section. Students are encouraged to use the Harvard Referencing Standard supplied by the Library. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action. Students may be required to undergo a viva (oral examination) if there is suspicion about the validity of their submitted work.

Signature: ……ADAEZE……

Date: ……7/12/2024……

PLEASE READ THE FOLLOWING INSTRUCTIONS:

1. Please attach a completed copy of this sheet to each project (including multiple copies).
2. Projects should be submitted to your Programme Coordinator.
3. **You must ensure that you retain a HARD COPY of ALL projects**, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer. Please do not bind projects or place in covers unless specifically requested.
4. You must ensure that all projects are submitted to your Programme Coordinator on or before the required submission date. **Late submissions will incur penalties.**
5. All projects must be submitted and passed in order to successfully complete the year. **Any project/assignment not submitted will be marked as a fail.**

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

AI Acknowledgement Supplement

[Insert Module Name]

[Insert Title of your assignment]

Your Name / Student Number	Course	Date

This section is a supplement to the main assignment, to be used if AI was used in any capacity in the creation of your assignment; if you have queries about how to do this, please contact your lecturer. For an example of how to fill these sections out, please click here.

AI Acknowledgment

This section acknowledges the AI tools that were utilized in the process of completing this assignment.

Tool Name	Brief Description	Link to tool

Description of AI Usage

This section provides a more detailed description of how the AI tools were used in the assignment. It includes information about the prompts given to the AI tool, the responses received, and how these responses were utilized or modified in the assignment. **One table should be used for each tool used**.

[Insert Tool Name]	
[Insert Description of use]	
[Insert Sample prompt]	[Insert Sample response]

Evidence of AI Usage

This section includes evidence of significant prompts and responses used or generated through the AI tool. It should provide a clear understanding of the extent to which the AI tool was used in the assignment. Evidence may be attached via screenshots or text.

Additional Evidence:

[Place evidence here]

Additional Evidence:

[Place evidence here]

Problem 1

I opened the NASA APOD page and entered my email to get API Key. I created a .env file where I saved my API-key and requirements.txt file that has all my dependencies. Then I ensured my .env connection on index.ipynb file by printing my API URL.

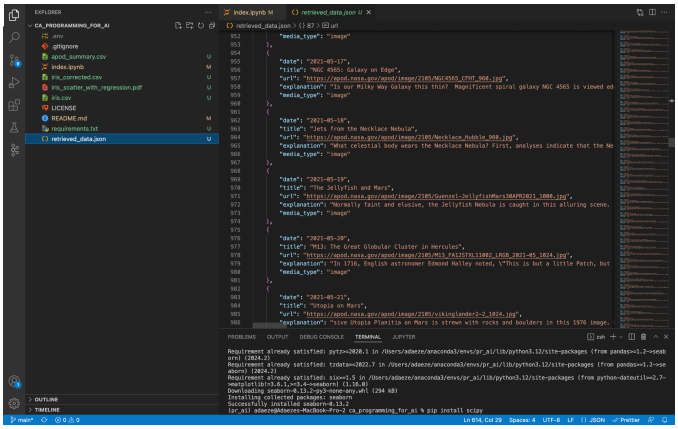
```
API URL is: https://api.nasa.gov
```

I created a function that gets the following information for 2020-01-01, Date of the picture, title of the picture, URL of the image or video, Explanation of the picture and Media type (image or video). To get this information I had to import REQUEST to enable the GET request.


```
{'date': '2020-01-01', 'title': 'Betelgeuse Imagined', 'url':
'https://apod.nasa.gov/apod/image/2001/BetelgeuseImagined_EsoCalcada_960.jpg', 'explanation': "Why is Betelgeuse fading?  No
one knows.  Betelgeuse, one of the brightest and most recognized stars in the night sky, is only half as bright as it used
to be only five months ago.  Such variability is likely just  normal behavior for this famously variable supergiant, but the
recent dimming has rekindled discussion on how long it may be before Betelgeuse does go supernova.  Known for its red color,
Betelgeuse is one of the few stars to be resolved by modern telescopes, although only barely.  The featured artist's
illustration imagines how Betelgeuse might look up close. Betelgeuse is thought to have a complex and tumultuous surface
that frequently throws impressive flares.  Were it to replace the Sun (not recommended), its surface would extend out near
the orbit of Jupiter, while gas plumes would bubble out past Neptune.  Since Betelgeuse is about 700 light years away, its
eventual supernova will not endanger life on Earth even though its brightness may rival that of a full Moon.  Astronomers --
both amateur and professional -- will surely continue to monitor Betelgeuse as this new decade unfolds.    Free
Presentation: APOD Editor to show best astronomy images of 2019 -- and the decade -- in NYC on January 3", 'media_type':
'image'}
```


I used try and expect to handle errors but I could not implement a specific error for invalid or wrong date format.

I generated data from Jan 1st to Dec 31 2020 while adding time.sleep 1 second to respect API rate limit. I used a while loop to iterate through the date range as I did not know the total count of data then tried to append the content of get_apod_data() but could not append JSON data as there is no format for appending JSON data. Retrieved_data.json file was created where collected data is saved and read from.



Problem 2

I wrote a function that reads the apod_data.json file and saves the data in a Python dictionary. I used try and expect to handle errors such as file not found and JSON errors.

```
Error: 'apod_data.json' file not found.
```

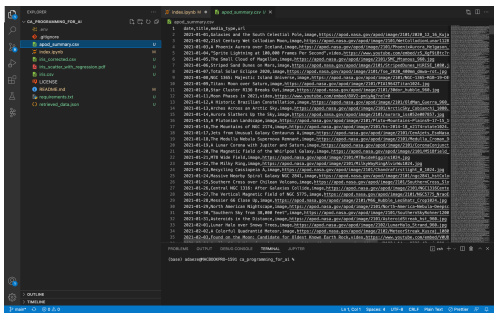
The Read_apod_data function returns only the date and title of each entry

```
{'date': '2021-01-01', 'title': 'Galaxies and the South Celestial Pole'}
{'date': '2021-01-02', 'title': '21st Century Wet Collodion Moon'}
{'date': '2021-01-03', 'title': 'A Phoenix Aurora over Iceland'}
{'date': '2021-01-04', 'title': 'Sprite Lightning at 100,000 Frames Per Second'
{'date': '2021-01-05', 'title': 'The Small Cloud of Magellan'}
{'date': '2021-01-06', 'title': 'Striped Sand Dunes on Mars'}
{'date': '2021-01-07', 'title': 'Total Solar Eclipse 2020'}
{'date': '2021-01-08', 'title': 'NGC 1365: Majestic Island Universe'}
{'date': '2021-01-09', 'title': 'Titan: Moon over Saturn'}
{'date': '2021-01-10', 'title': 'Star Cluster R136 Breaks Out'}
{'date': '2021-01-11', 'title': 'Moon Phases in 2021'}
{'date': '2021-01-12', 'title': 'A Historic Brazilian Constellation'}
{'date': '2021-01-13', 'title': 'Arches Across an Arctic Sky'}
{'date': '2021-01-14', 'title': 'Aurora Slathers Up the Sky'}
{'date': '2021-01-15', 'title': 'A Plutonian Landscape'}
{'date': '2021-01-16', 'title': 'The Mountains of NGC 2174'}
```

To get the total count for images, videos and longest text, I created a new function, initialized image and video count to 0 then looped through the data checking for where media type is either image or video then increasing their count. For the longest text I used Python library max() and len() to get the entry with the largest date.

```
Total count of Images: 328
Total count of Videos: 37
Entry with the longest explanation:
{
    "date": "2021-04-16",
    "title": "The Doubly Warped World of Binary Black Holes",
    "url": "https://www.youtube.com/embed/rQcKIN9vj3U?rel=0",
    "explanation": "Light rays from accretion disks around a pair of orbiting supermassive black holes make their way
through the warped space-time produced by extreme gravity in this stunning computer visualization. The simulated accretion
disks have been given different false color schemes, red for the disk surrounding a 200-million-solar-mass black hole, and
blue for the disk surrounding a 100-million-solar-mass black hole. That makes it easier to track the light sources, but the
choice also reflects reality. Hotter gas gives off light closer to the blue end of the spectrum and material orbiting
smaller black holes experiences stronger gravitational effects that produce higher temperatures. For these masses, both
accretion disks would actually emit most of their light in the ultraviolet though. In the video, distorted secondary images
of the blue black hole, which show the red black hole's view of its partner, can be found within the tangled skein of the
red disk warped by the gravity of the blue black hole in the foreground. Because we're seeing red's view of blue while also
seeing blue directly, the images allow us to see both sides of blue at the same time. Red and blue light originating from
both black holes can be seen in the innermost ring of light, called the photon ring, near their event horizons. Astronomers
expect that in the not-too-distant future they\u2019ll be able to detect gravitational waves, ripples in space-time,
produced when two supermassive black holes in a system much like the one simulated here spiral together and merge.",
    "media_type": "video"
}
```

I used a function to extract and write data to apod_summary CSV. The CSV file contains appropriate column headers. New data is appended to existing data, see output below



Problem 3

I imported numpy and then created 20 rows and 5 columns with .random() ranging from 10 to 100, which I named data.

```
array([[13, 85, 52, 47, 98],
       [85, 62, 91, 79, 81],
       [85, 41, 63, 55, 45],
       [39, 49, 90, 21, 63],
       [95, 84, 78, 18, 84],
       [57, 90, 29, 64, 35],
       [24, 96, 39, 45, 40],
       [37, 75, 92, 73, 11],
       [83, 15, 98, 21, 29],
       [26, 54, 99, 95, 92],
       [32, 91, 65, 37, 21],
       [12, 44, 67, 36, 72],
       [76, 67, 40, 29, 59],
       [85, 52, 49, 63, 26],
       [46, 71, 58, 29, 16],
       [16, 66, 89, 23, 61],
       [80, 66, 29, 89, 94],
       [52, 12, 43, 12, 13],
       [61, 53, 29, 56, 32],
       [94, 71, 58, 62, 35]])
```

```
# Sums each row to check if they are even or odd.
np.sum(data, axis=1)
```

Summed all data to check if they were odd or even, since they were not all even I had to loop through each row, rows that summed up to even were displayed and for the odd rows, I subtracted 1 from the first data on that row to make it even. See output below

```
260
192
270
334
276
236
186
194
204
228
306
332
304
162
308
270
244
306
252
216
```

Multiples of 5 are numbers when divided by 5 has 0 reminder, so to check if the total is a multiple of 5 I had to sum the whole array and divide it by 5, if it returns 0 it is a multiple of 5 and the total is displayed but if it is not the reminder is subtracted from one element in the row to make it a multiple of 5.

For this, an AND operator is used to ensure the values displayed are divisible by both 3 and 5.

The mean of the array is gotten, numpy has a function for this. Then a check is made using an if statement to check for numbers greater than 75 and the numbers found were replaced by mean. Using numpy functions I calculated the mean, Standard Deviation, median and Variance for each column

```
Mean of the array: 46.36
Standard Deviation of the array: 17.010890629241022
Median of the array: 55.0
Variance for each column: [314.84   256.3275 129.11   314.86
```

Problem 4

I used iris data from Moodle a copy has been included in the project and all necessary libraries have been imported. There are 150 rows and 5 columns in the dataset. There are 4 numerical columns and a Nominal Categorical Feature (Species), see below for the names of the columns

```
Index(['Sepal.Length', 'Sepal.Width', 'Petal.Length', 'Petal.Width',
       'Species'],
```

There are 3 species of flower

```
array(['setosa', 'versicolor', 'virginica'], dtype=object)
```

The 2 rows with incorrect data were fixed

```
# Checking fix.
updatedRows = irisData.iloc[[35,38]]
print(updatedRows)
```

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
35	4.9	3.1	1.5	0.2	setosa
38	4.9	3.6	1.4	0.1	setosa

Two extra columns were added and they are derived from the ratio of the width and length of existing features

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species	Petal.Ratio	Sepal.Ratio
0	5.1	3.5	1.4	0.2	setosa	7.000000	1.457143
1	4.9	3.0	1.4	0.2	setosa	7.000000	1.633333
2	4.7	3.2	1.3	0.2	setosa	6.500000	1.468750
3	4.6	3.1	1.5	0.2	setosa	7.500000	1.483871
4	5.0	3.6	1.4	0.2	setosa	7.000000	1.388889

This new dataset is saved to CSV

4.4 The two pairs of features that have the highest positive correlation are **petal width and petal length with 0.962786 as they are close to 1** and the highest negative correlation are **sepal length and width with -0.119003 this indicates a weak negative linear**

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Petal.Ratio	Sepal.Ratio
Sepal.Length	1.000000	-0.119003	0.871693	0.817809	-0.564533	0.728583
Sepal.Width	-0.119003	1.000000	-0.429425	-0.365543	0.335022	-0.750309
Petal.Length	0.871693	-0.429425	1.000000	0.962786	-0.684523	0.836973
Petal.Width	0.817809	-0.365543	0.962786	1.000000	-0.733923	0.752458
Petal.Ratio	-0.564533	0.335022	-0.684523	-0.733923	1.000000	-0.589552
Sepal.Ratio	0.728583	-0.750309	0.836973	0.752458	-0.589552	1.000000

From these statistics, I can infer that petal width and length are highly correlated and have a strong linear relationship meaning as petal length increases, petal width increases as well. While sepal width and length are weakly related meaning a change in sepal width might not lead to a change in sepal length.