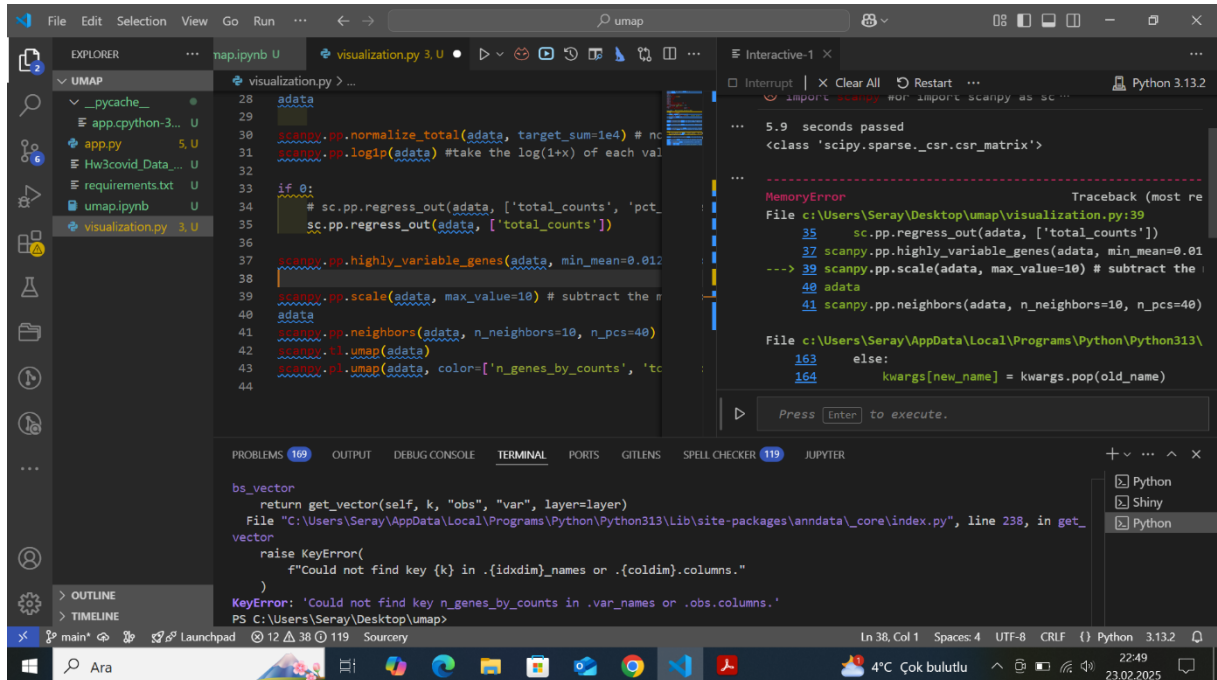


ERRORS

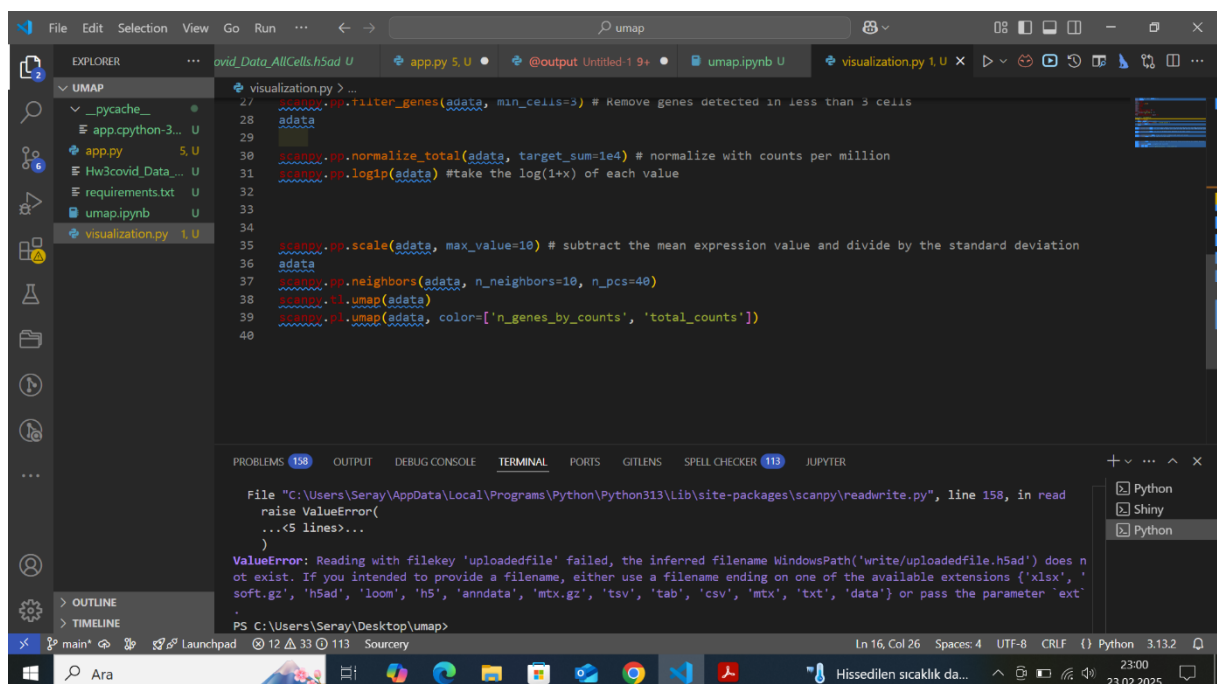
There have been some challenges in this Project. Here are 5 of them that I've understood.

1. Memory Error



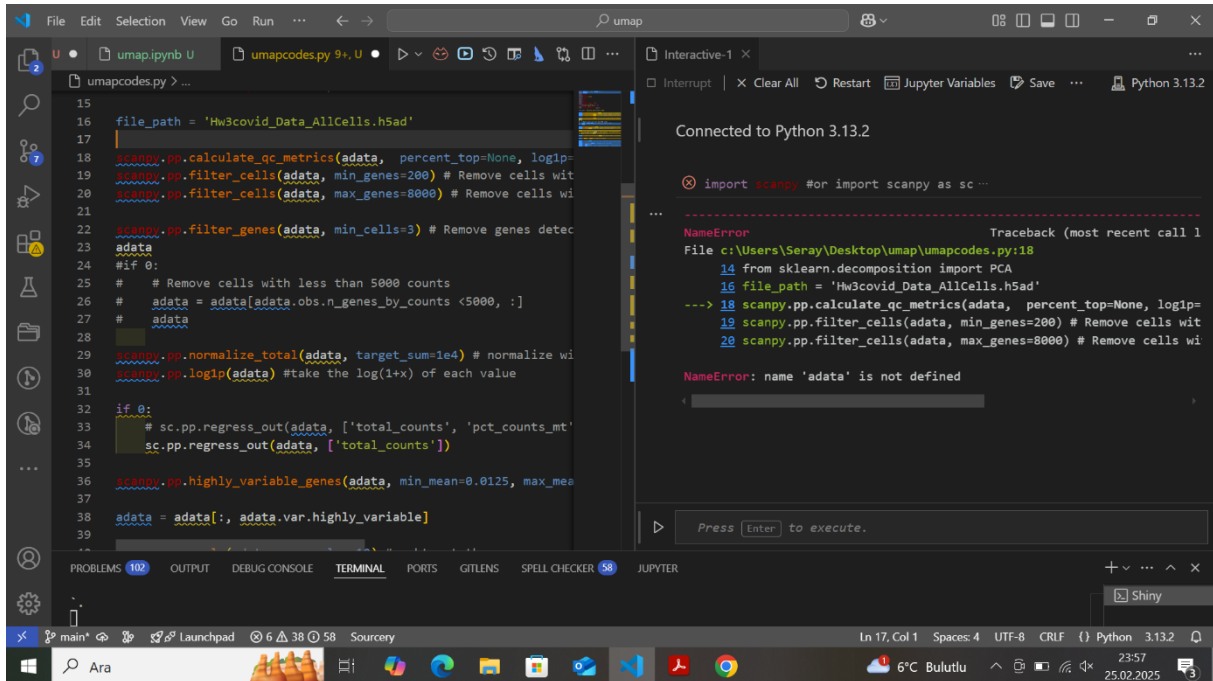
I've faced this error because of the fact that I've deleted essential codes being used in the data cleaning and clustering part. In this case we have a very large dataset which needs to be properly handled if not, it could result in errors such as these. What I did was to write the needed codes so that we can work with only the needed part.

2. ValueError



Here, I've gotten a value error because I've created a variable with the name uploadedfile but infact didn't add any values to it. What I tried to do was creating an empty variable. I couldn't fix the problem here so I deleted the variable and continued with the already existing ones so that they connected and the variable was no longer unmeaningful.

3. NameError



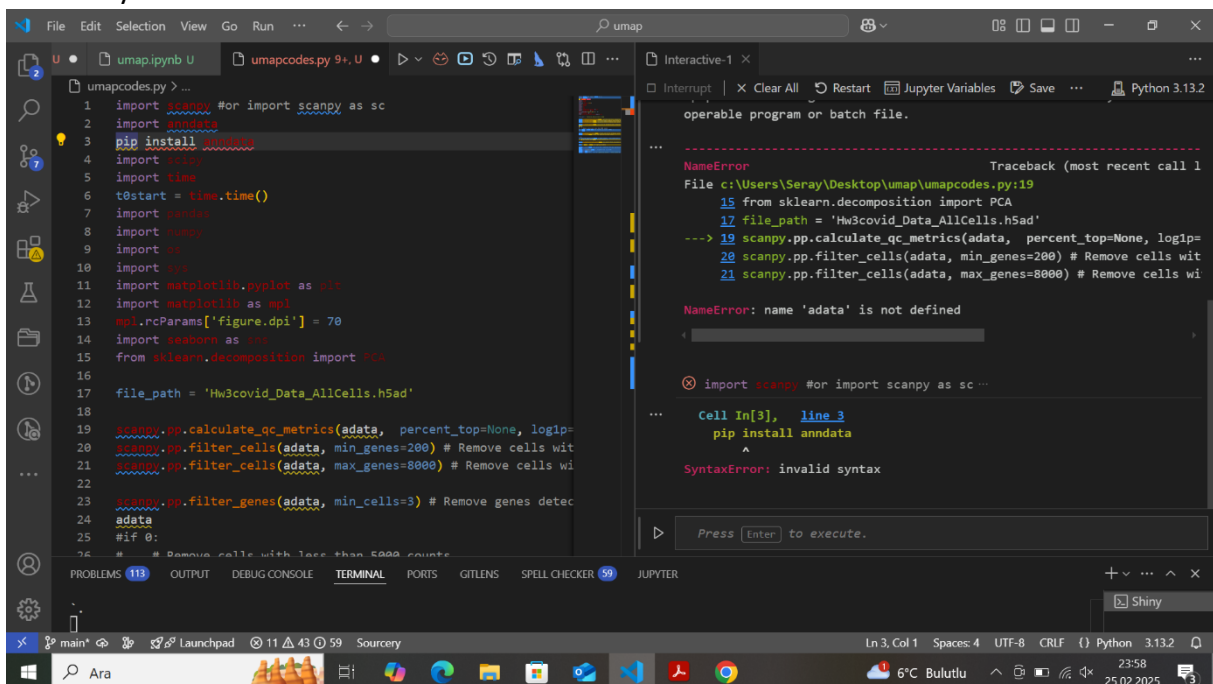
```
15 file_path = 'Hw3covid_Data_AllCells.h5ad'
16
17
18 scanpy.pp.calculate_qc_metrics(adata, percent_top=None, log1p=
19 scanpy.pp.filter_cells(adata, min_genes=200) # Remove cells wit
20 scanpy.pp.filter_cells(adata, max_genes=8000) # Remove cells wi
21
22 scanpy.pp.filter_genes(adata, min_cells=3) # Remove genes detec
23 adata
24 #if 0:
25 # # Remove cells with less than 5000 counts
26 # adata = adata[adata.obs.n_genes_by_counts < 5000, :]
27 # adata
28
29 scanpy.pp.normalize_total(adata, target_sum=1e4) # normalize wi
30 scanpy.pp.log1p(adata) #take the log(1+x) of each value
31
32 if 0:
33 # sc.pp.regress_out(adata, ['total_counts', 'pct_counts_mt'
34 sc.pp.regress_out(adata, ['total_counts'])
35
36 scanpy.pp.highly_variable_genes(adata, min_mean=0.0125, max_mea
37
38 adata = adata[:, adata.var.highly_variable]
39
```

Traceback (most recent call 1)
File c:\Users\Seray\Desktop\umap\umapcodes.py:18
14 from sklearn.decomposition import PCA
16 file_path = 'Hw3covid_Data_AllCells.h5ad'
----> 18 scanpy.pp.calculate_qc_metrics(adata, percent_top=None, log1p=
19 scanpy.pp.filter_cells(adata, min_genes=200) # Remove cells wit
20 scanpy.pp.filter_cells(adata, max_genes=8000) # Remove cells wi

NameError: name 'adata' is not defined

I've solved this problem by gathering the codes in a function. When I used def adata was no longer unknown to the code.

4. Syntax Error



```
1 import scanpy #or import scanpy as sc
2 import anndata
3 pip install anndata
4 import time
5 import time
6 t0start = time.time()
7 import pandas
8 import numpy
9 import os
10 import sys
11 import matplotlib.pyplot as plt
12 import matplotlib as mpl
13 mpl.rcParams['figure.dpi'] = 70
14 import seaborn as sns
15 from sklearn.decomposition import PCA
16
17 file_path = 'Hw3covid_Data_AllCells.h5ad'
18
19 scanpy.pp.calculate_qc_metrics(adata, percent_top=None, log1p=
20 scanpy.pp.filter_cells(adata, min_genes=200) # Remove cells wit
21 scanpy.pp.filter_cells(adata, max_genes=8000) # Remove cells wi
22
23 scanpy.pp.filter_genes(adata, min_cells=3) # Remove genes detec
24 adata
25 #if 0:
26 # # Remove cells with less than 5000 counts
```

operable program or batch file.
Traceback (most recent call 1)
File c:\Users\Seray\Desktop\umap\umapcodes.py:19
15 from sklearn.decomposition import PCA
17 file_path = 'Hw3covid_Data_AllCells.h5ad'
----> 19 scanpy.pp.calculate_qc_metrics(adata, percent_top=None, log1p=
20 scanpy.pp.filter_cells(adata, min_genes=200) # Remove cells wit
21 scanpy.pp.filter_cells(adata, max_genes=8000) # Remove cells wi

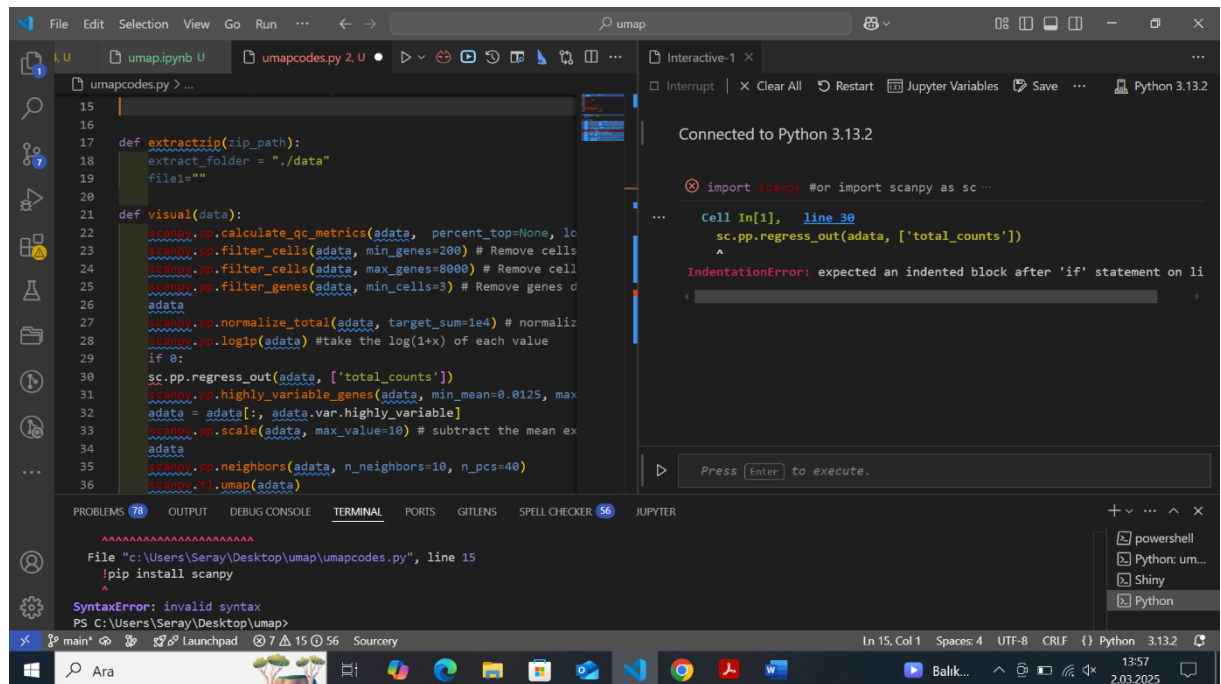
NameError: name 'adata' is not defined

import scanpy #or import scanpy as sc

Cell In[3], line 3
pip install anndata
^
SyntaxError: invalid syntax

I'm still not really sure why it kept giving syntax error when I tried to pip install anndata because I tried both typing it with ! and without. Because of this I'm thinking the main error was connected to gathering the codes in a function. Also maybe it gave this error because anndata was already installed. I deleted this unnecessary code.

5. IdentitationError



```
15
16
17 def extractzip(zip_path):
18     extract_folder = "./data"
19     file1=""
20
21 def visual(data):
22     scanpy.pp.calculate_qc_metrics(adata, percent_top=None, log1p=False, inplace=True)
23     scanpy.pp.filter_cells(adata, min_genes=200) # Remove cells
24     scanpy.pp.filter_cells(adata, max_genes=8000) # Remove cell
25     scanpy.pp.filter_genes(adata, min_cells=3) # Remove genes d
26     adata
27     scanpy.pp.normalize_total(adata, target_sum=1e4) # normaliz
28     scanpy.pp.log1p(adata) #take the log(1+x) of each value
29     if 0:
30         sc.pp.regress_out(adata, ['total_counts'])
31         scanpy.pp.highly_variable_genes(adata, min_mean=0.0125, max
32         adata = adata[:, adata.var.highly_variable]
33         scanpy.pp.scale(adata, max_value=10) # subtract the mean ex
34         adata
35         scanpy.pp.neighbors(adata, n_neighbors=10, n_pcs=40)
36         scanpy.tl.umap(adata)
```

Problems (78) OUTPUT DEBUG CONSOLE TERMINAL PORTS GIT LENS SPELL CHECKER (56) JUPYTER

File "c:\Users\Seray\Desktop\umap\umapcodes.py", line 15
!pip install scanpy
SyntaxError: invalid syntax
PS C:\Users\Seray\Desktop\umap>

This happened because the code was not inside the if function. But this was not that necessary so I deleted the code instead.