MP1

ECE 471
Fall 2024

Zach Larson (zlarson2), Máximo Rojas (mgr9), Matthew Paul (mjpaul3)

Group 9

Task 1

1. Import all the .csv data (including 4 weather conditions: "clear-night", "clearsunset", "clear-noon", "rain-noon") of the ghost_cutin scene into Jupyter Notebook. List the first 5 rows of the ctl.csv, cvip.csv, and traj.csv with the clear-night weather condition.

From the simulation with Carla, our team obtained data for 6 EV driving scenarios with different weather conditions. For each scenario, the data was split into 3 different files and consisted of the following features:

- Timestamp (ts)
- ID of the AV (agent_ID)
- Acceleration of the AV (throttle)
- Direction of the AV (steer)
- Braking of the AV (brake)
- Distance of AV to NPC Actor (cvip)
- Horizontal location of AV (x)
- Moving direction of AV (y)
- Speed of AV along y-axis (v)

Clear night: _ctl.csv

	ts	agent_id	throttle	steer	brake
0	34097	0	0.9	-0.013192	0.0
1	34098	0	0.9	-0.003670	0.0
2	34099	0	0.9	-0.004190	0.0
3	34100	0	0.9	-0.003286	0.0
4	34101	0	0.9	0.000093	0.0

Clear night: _cvip.csv

	ts	agent_id	cvip	cvip_x	cvip_y	cvip_z
0	34097	0	500.491189	198.767441	-95.832657	-499.819366
1	34098	0	5.595580	195.567444	-90.832657	0.100000
2	34099	0	5.592365	195.567444	-90.832657	0.095407
3	34100	0	5.589578	195.567444	-90.832657	0.084691
4	34101	0	5.587154	195.567444	-90.832657	0.069311

Clear night: _traj.csv

	ts	agent_id	x	y	z	v
0	34097	0	192.362411	-86.26268	0.539326	0.0
1	34098	0	192.362411	-86.26268	0.491906	0.0
2	34099	0	192.362411	-86.26268	0.438374	0.0
3	34100	0	192.362411	-86.26268	0.378732	0.0
4	34101	0	192.362411	-86.26268	0.312981	0.0

Task 1

2. Summarize the following information for each weather condition ("clear-night", "clear-sunset", "clear-noon", "rain-noon"): (2 points)
   a. The duration of the scene.
   b. Mean and standard deviation of the values of the features (“throttle”, “steer”, “brake”, “cvip”, “x”, “y”, “v”). Round your results to 3 decimal place and save them in a table, with the weather conditions as columns, and the features as rows (hint: you can store the table in a data frame).

Means Table

	Clear Night	Clear Sunset	Clear Noon	Rain Noon	Haze Noon	Haze Sunset
throttle	0.633	0.610	0.609	0.644	0.601	0.601
steer	0.004	0.000	0.001	-0.000	0.001	0.001
brake	0.116	0.057	0.061	0.047	0.055	0.059
cvip	32.185	19.488	18.826	5.828	19.118	19.730
x	191.313	192.947	192.921	192.631	192.941	192.944
y	-32.170	-31.068	-31.347	-63.818	-31.725	-31.407
v	6.223	6.893	6.938	6.358	6.927	6.877

Standard Deviations Table

	Clear Night	Clear Sunset	Clear Noon	Rain Noon	Haze Noon	Haze Sunset
throttle	0.345	0.281	0.282	0.273	0.298	0.304
steer	0.048	0.005	0.005	0.003	0.005	0.005
brake	0.320	0.232	0.240	0.213	0.227	0.236
cvip	27.841	17.432	16.721	1.260	16.415	17.262
x	1.171	0.400	0.404	0.253	0.411	0.410
y	39.071	40.427	40.487	21.141	40.711	40.589
v	2.877	3.307	3.277	3.856	3.043	3.055

Scene Durations

	Clear Night	Clear Sunset	Clear Noon	Rain Noon	Haze Noon	Haze Sunset
Duration (ms)	838	756	750	400	751	757

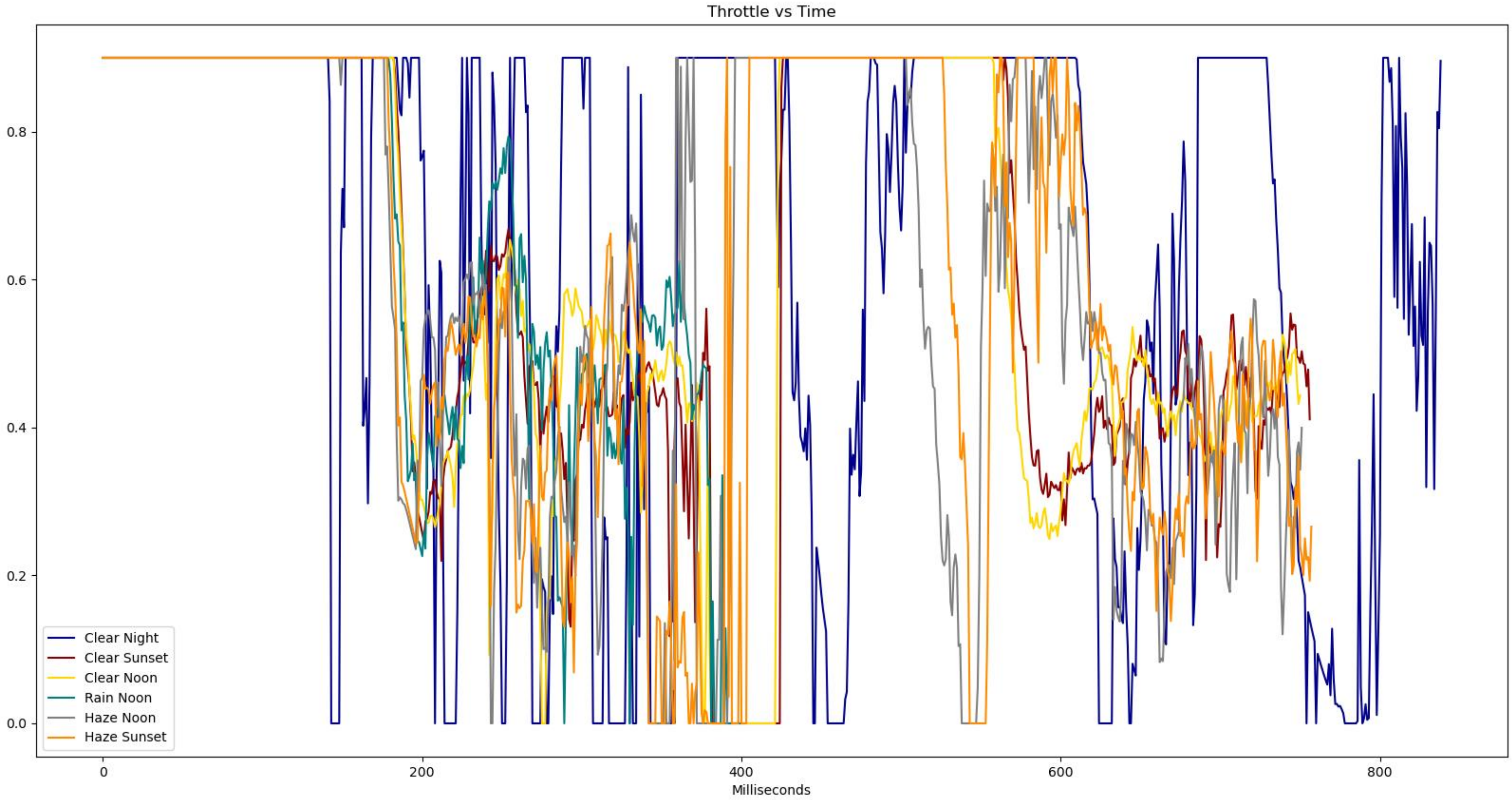
Task 1

3. Visualize the campaign results of the ghost_cutin scene for each weather condition. Plot the throttle values (y-axis) of the agent vs time (x-aixs). Please plot all the weather conditions in one figure, and repeat the same step for all other features as well (“steer”, “brake”, “cvip”, “x”, “y”, “v”).

PLOTS ON FOLLOWING SLIDES

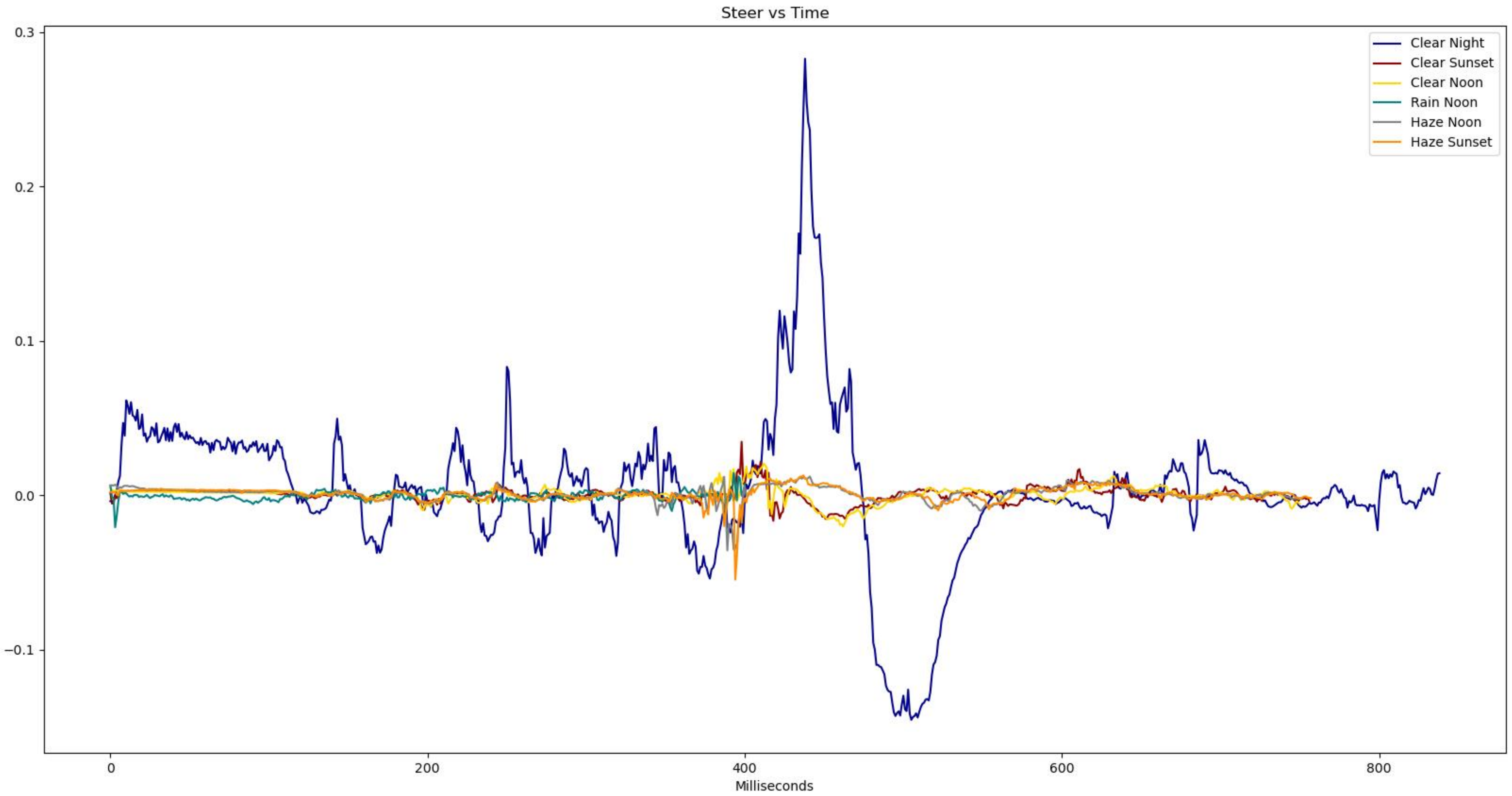
Task 1

Campaign Result Visualization



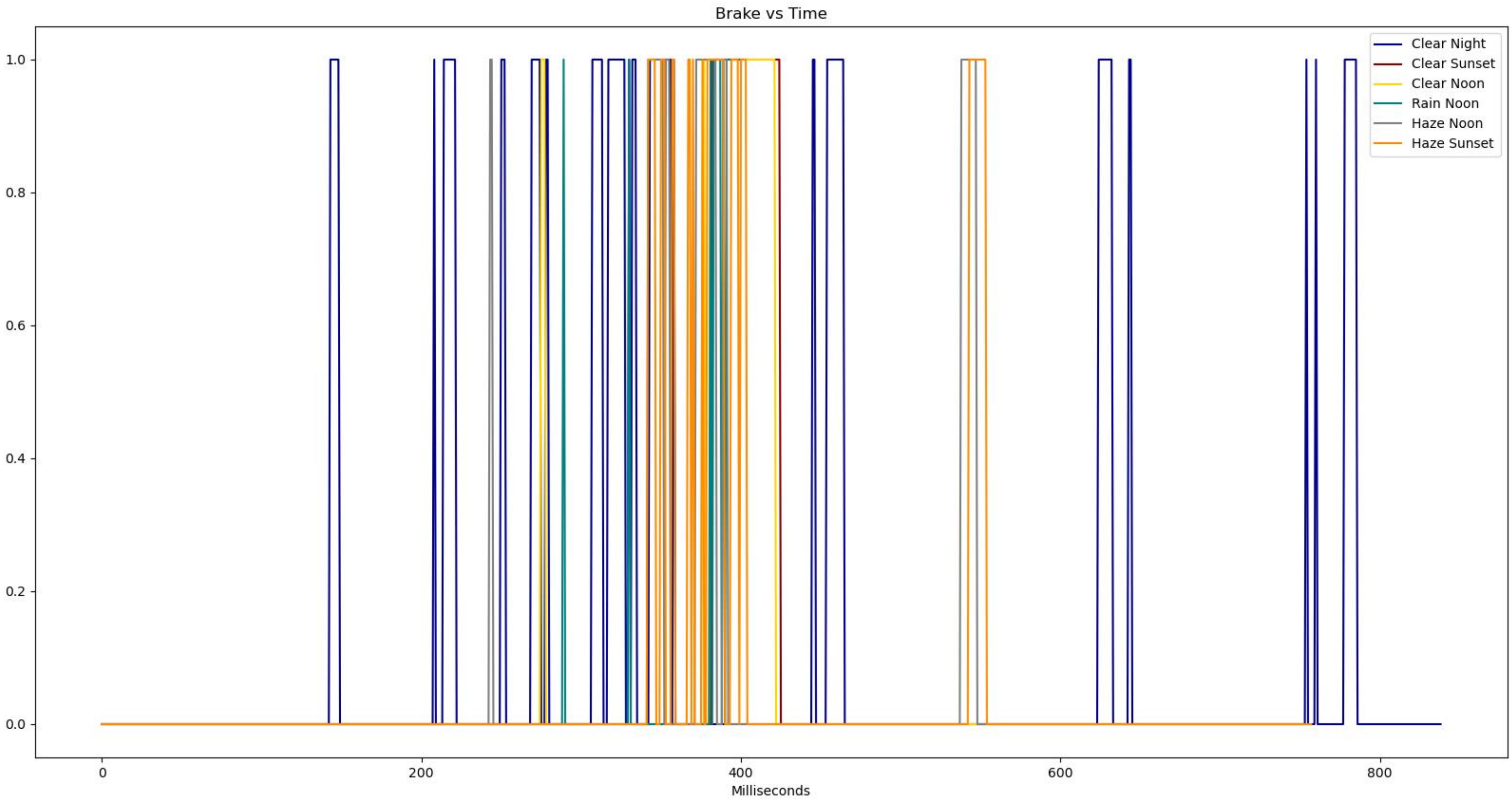
Task 1

Campaign Result Visualization



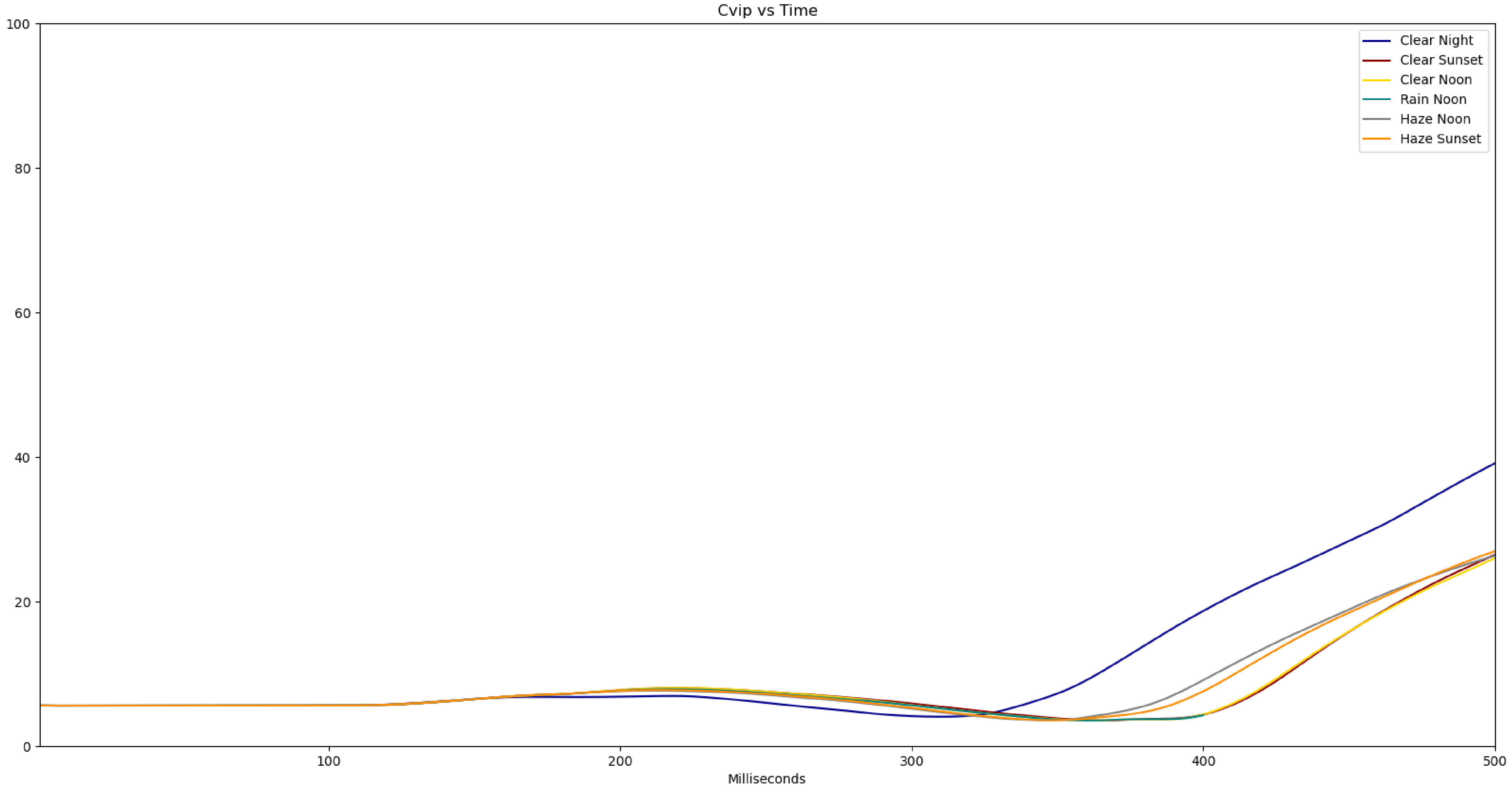
Task 1

Campaign Result Visualization



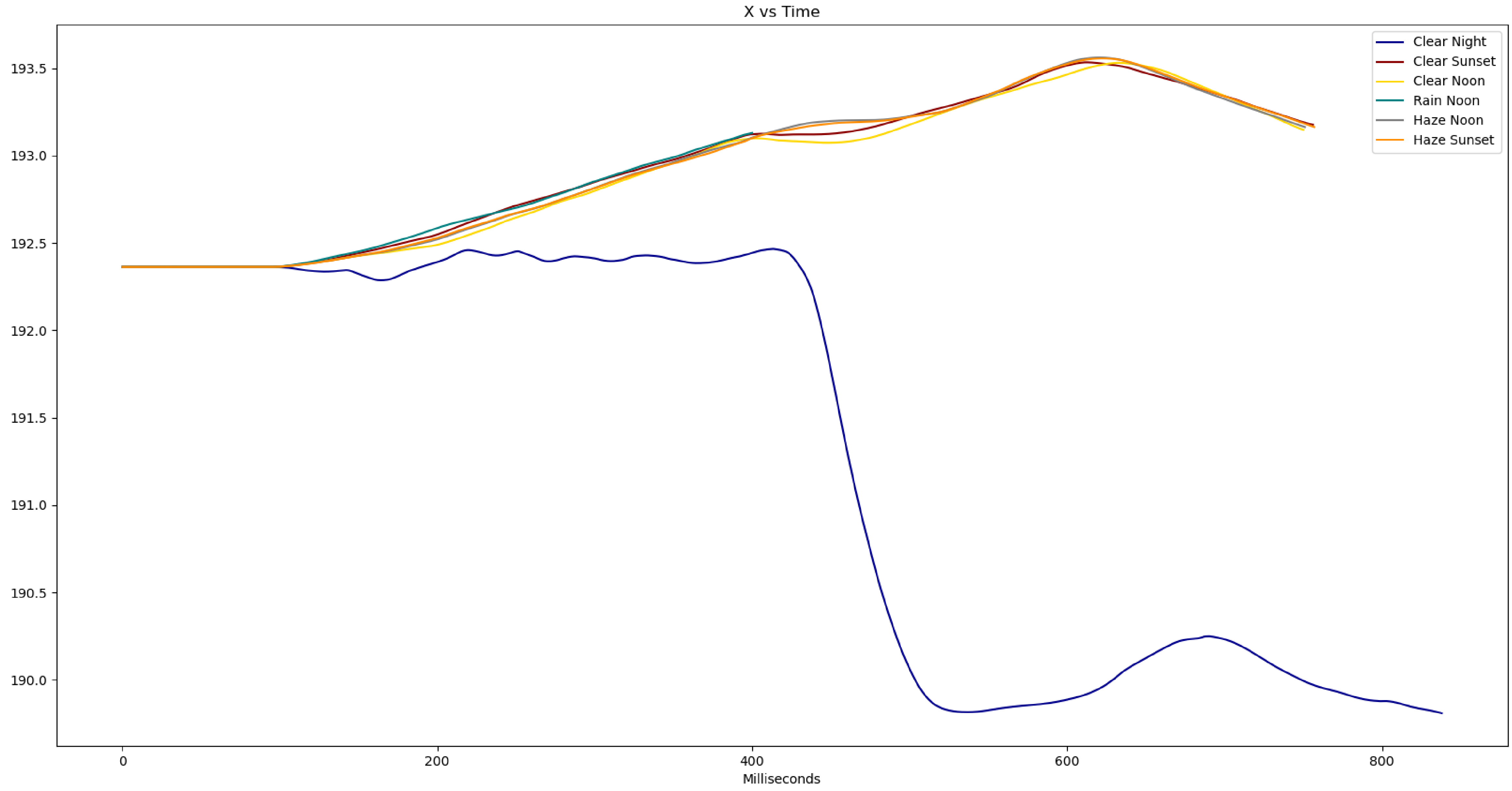
Task 1

Campaign Result Visualization



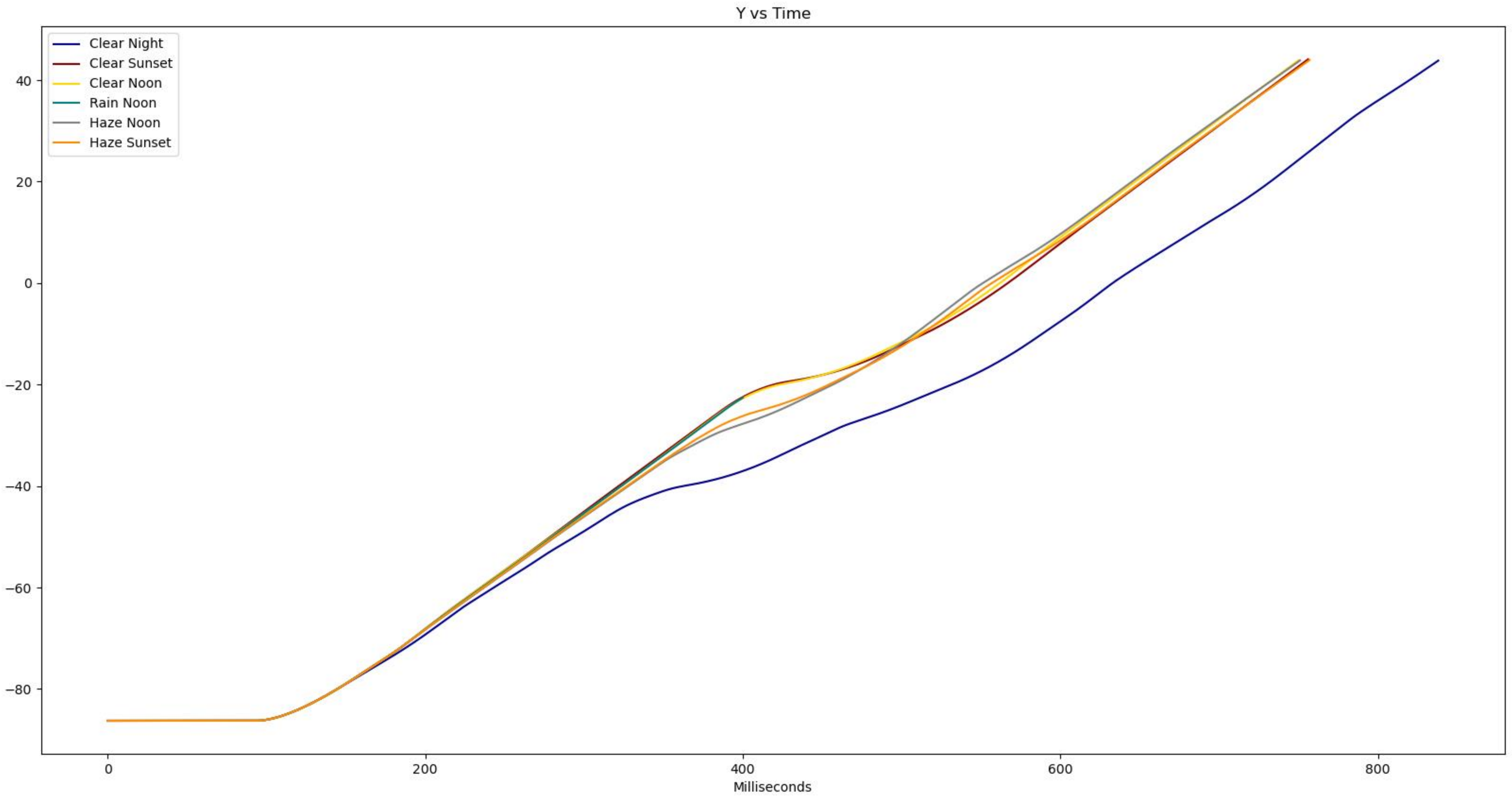
Task 1

Campaign Result Visualization



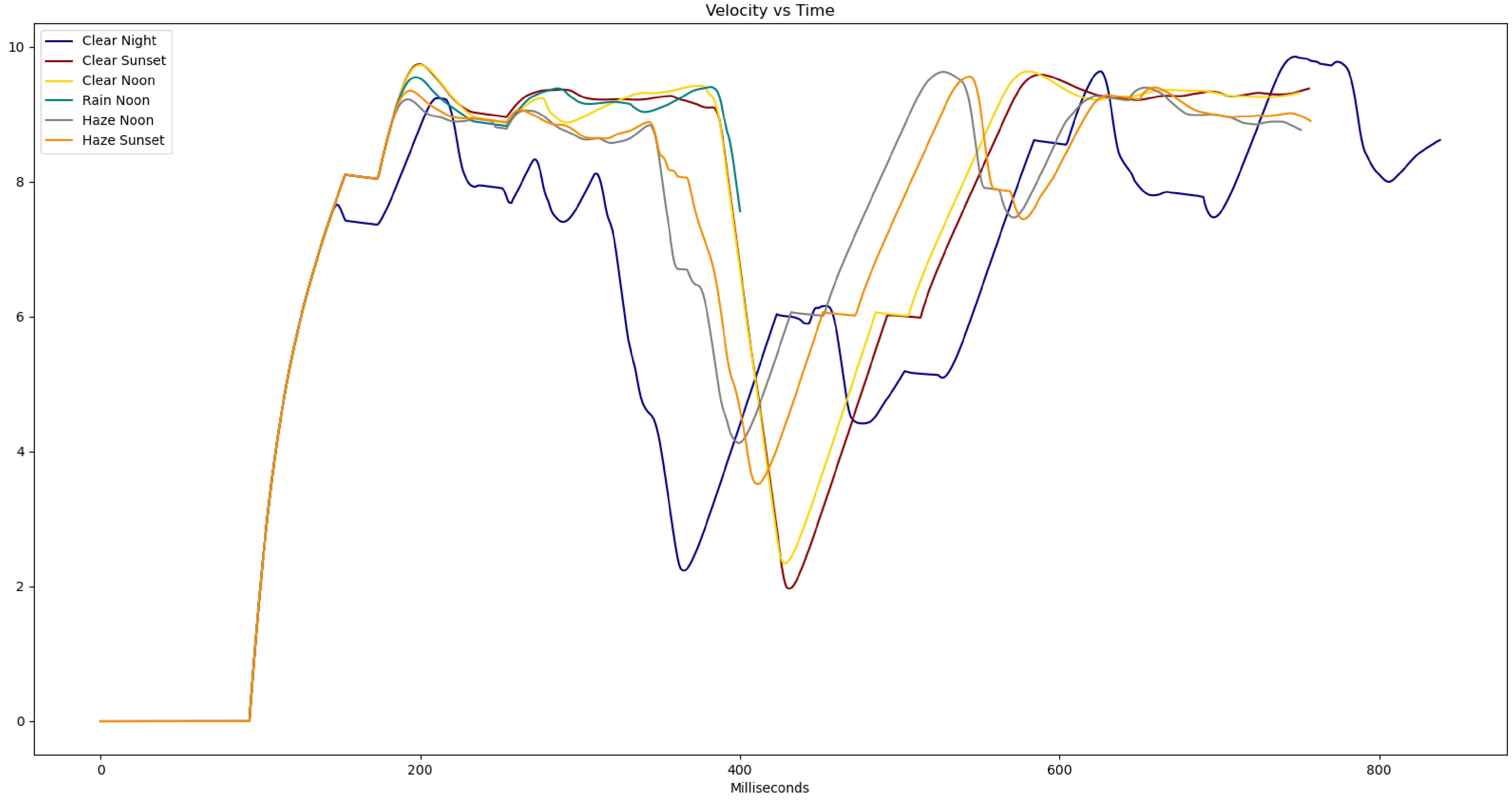
Task 1

Campaign Result Visualization



Task 1

Campaign Result Visualization



Task 1

4. Based on your intuition and life experience, which of the features do you think will change during an accident? How will the feature(s) change? By looking at the plots you generated in Task 1.3, combined with your reasoning (without looking at 'route_highway.txt'), which weather condition(s) has an accident?

- throttle: would expect to decrease or go to zero
- velocity: would expect to suddenly drop
- steer: direction may or may not change
- brake: may or may not change depending on the abruptness of the crash. Most likely sudden brake before impact if noticed
- x: correlated to steering, would expect a sudden change when merging to another lane
- y: expect a change in slope when accelerating or braking
- cvip: would expect it to go to zero before the crash

Based on the throttle, steering, and change in x-location in the clear night scenario, our group predicts a crash in that condition. No other EV in each of the different scenarios appears to have such volatile driving characteristics as compared to in this scenario.

Task 2

1. Suppose each simulation run has a result of accident/non-accident, calculate the probability of accident (counts, marginal probability). Hint: for each run, the collision results are stored in 'route_highway.txt'. You can check the accident status by looking at the 'status' field under the 'record' section ('Completed' means no accident; 'Failed' means an accident has occurred).
2. By looking at the completion records and the plots you generated in Task 1, under which weather condition(s) did the accident happen? Does that match your guess in Task 1? When did the accident happen during those simulation runs? Why do you think the accident happened at that instance? Discuss each accident case separately.

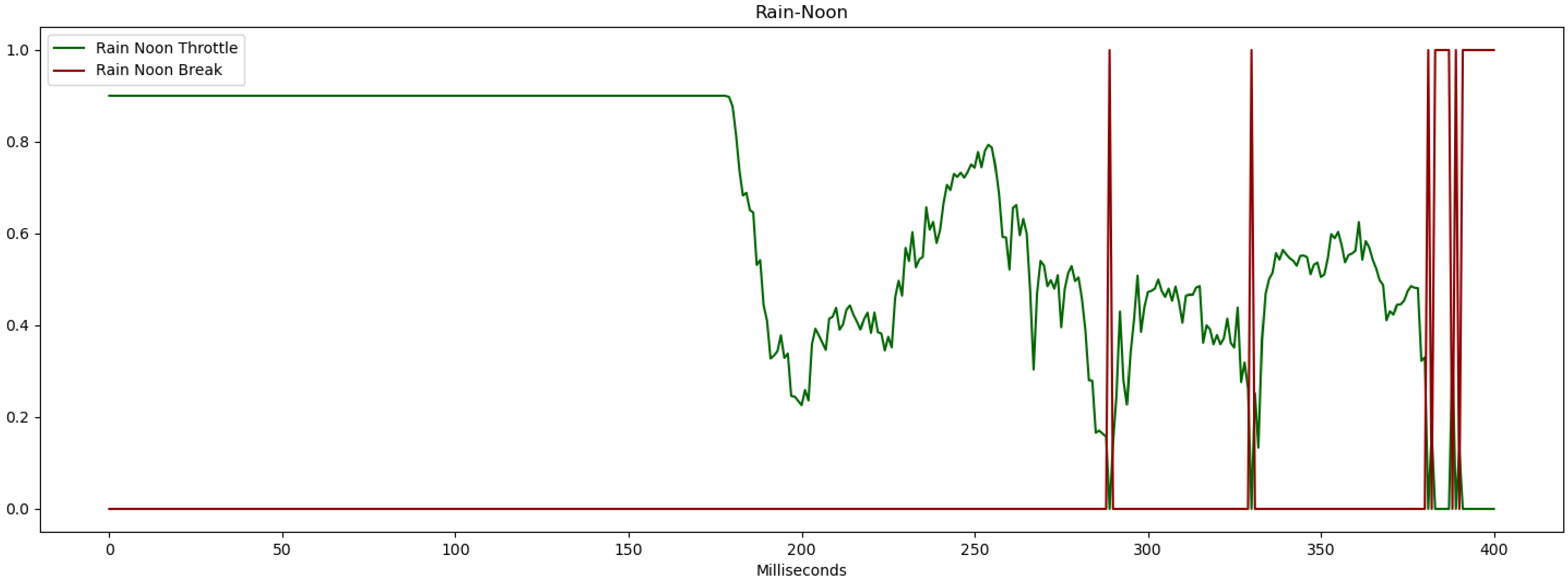
The probability of accident is 1/6.

Our data includes 6 simulations in distinct weather conditions. There was only one recorded crash, which took place in the rain-noon simulation. Therefore, without knowing the weather conditions, the probability of a crash in any given condition would be 1/6.

The crash occurred in the rain-noon weather condition, which does not match the prediction from Task 1. It took place slightly after 400 milliseconds into the simulation. The graphs give no obvious insight into why the crash occurred exactly at that time. Some minor observations from the plots include the following:

- Insufficient breaking: the graph reveals that the vehicle did not break consistently during the time right before the crash. Additionally, it began breaking later than in the other cases. This lines up with what we viewed in the simulation through Carla.
- Throttle: From the throttle vs time graph, we can see the vehicle accelerating immediately preceding the crash.
- Cvip: The EV remained consistently close to the risky NPC actor throughout the duration of the simulation.

Task 2



Task 2

3. Accidents are caused by abnormal AV behavior. However, there are other adverse driving conditions when there are abnormal AV behaviors while no accident occurs. From the plots you generated in Task 1.3, do you observe any other abnormal behavior? If so, what do you think is (are) the cause(s) of this behavior?

4. In this question, we explore differences between abnormal and normal runs. Complete the following questions (6 points)

    a. We study the following features: “brake”, “steer”, “v”, “y”, “x”, “cvip”, “throttle”. Plot the distribution of each feature for the abnormal runs (including the accident runs) vs normal runs. Treat the values at each time point as an independent individual sample and generate the density plot of the distribution. Describe the difference between the “steer” distribution for normal and abnormal runs.

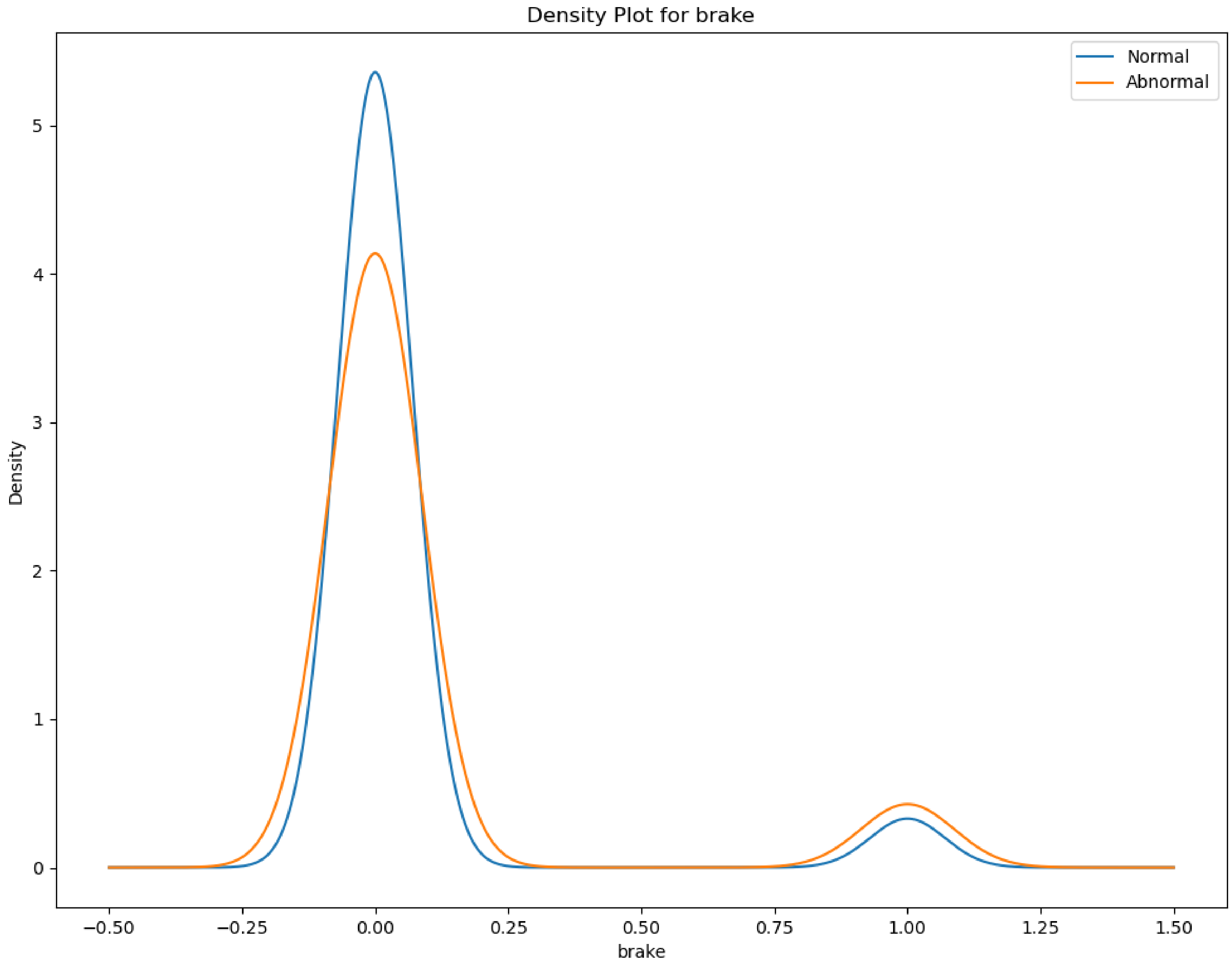
There was one other weather condition in which the EV displayed atypical behavior. As mentioned previously, this was the clear night scenario. Some of the atypical behavior included:

- Consistently lower speeds
- Lane change
- Jittery steering, even before lane change
- Throttle usage very jumpy, typically either fully on or fully off
- Maintained maximum distance from NPC actor

We believe many of these observations are due to EV nighttime programmed behavior probably being more risk-adverse and aware of surroundings.

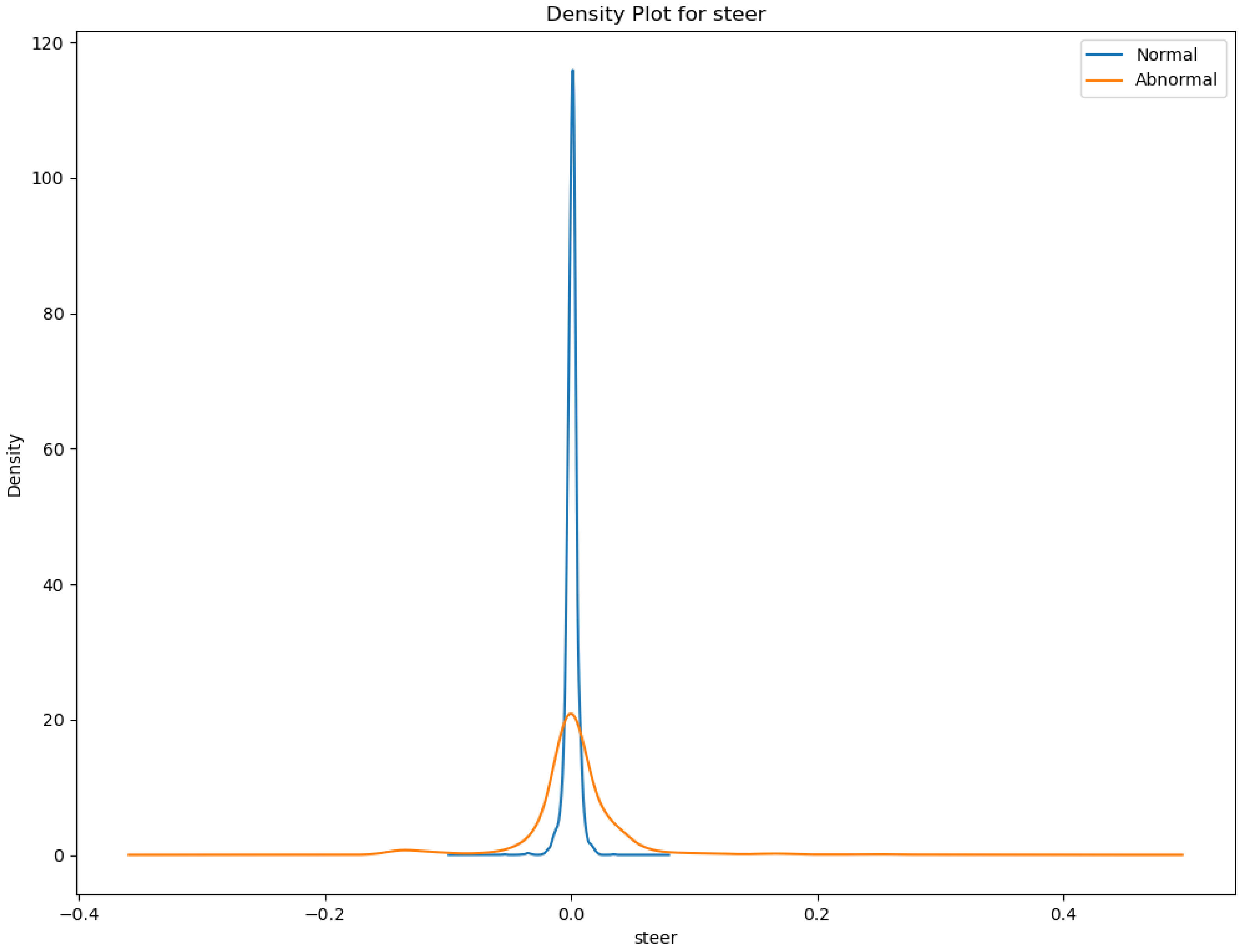
Task 2

Distribution of the Features: Normal vs Abnormal



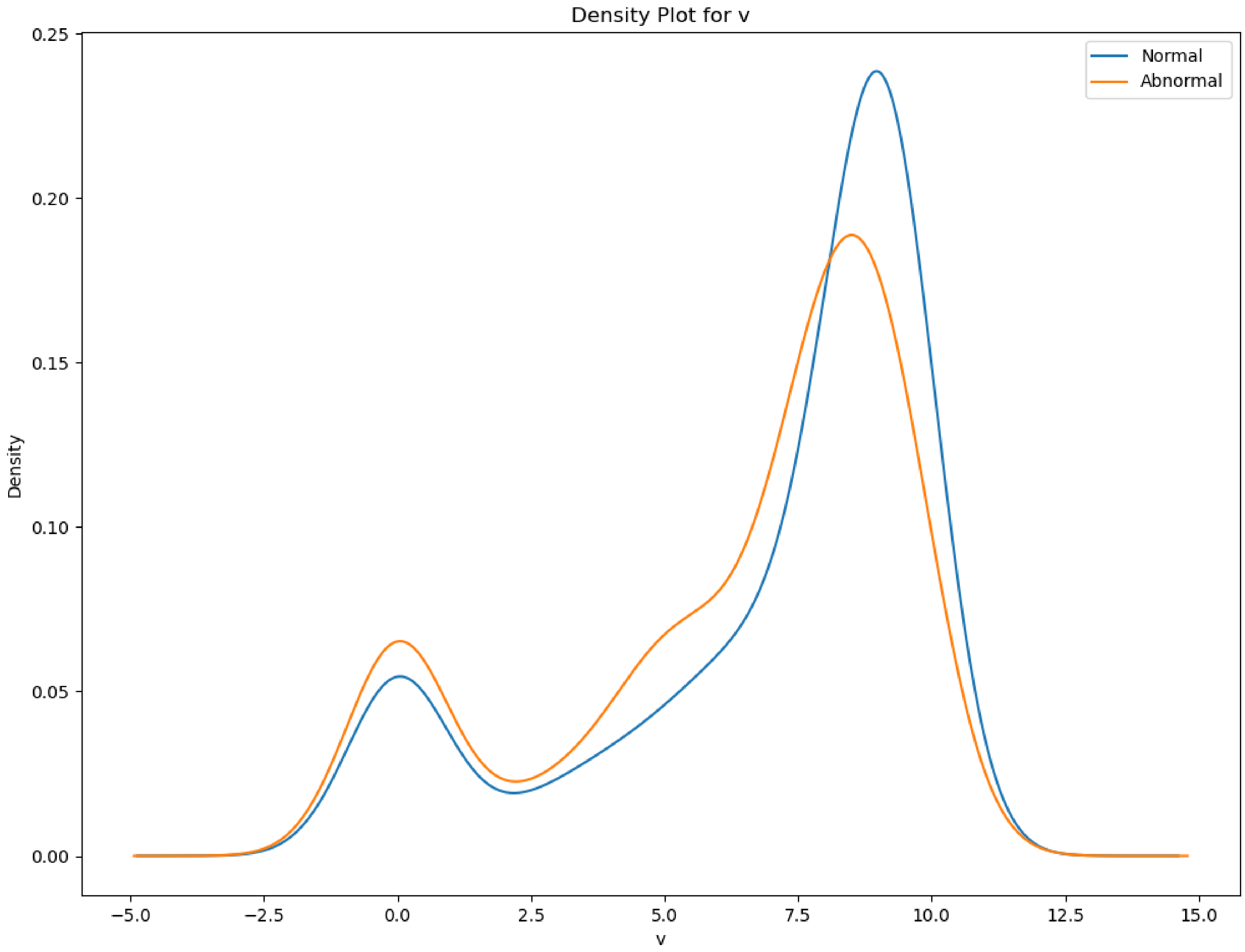
Task 2

Distribution of the Features: Normal vs Abnormal



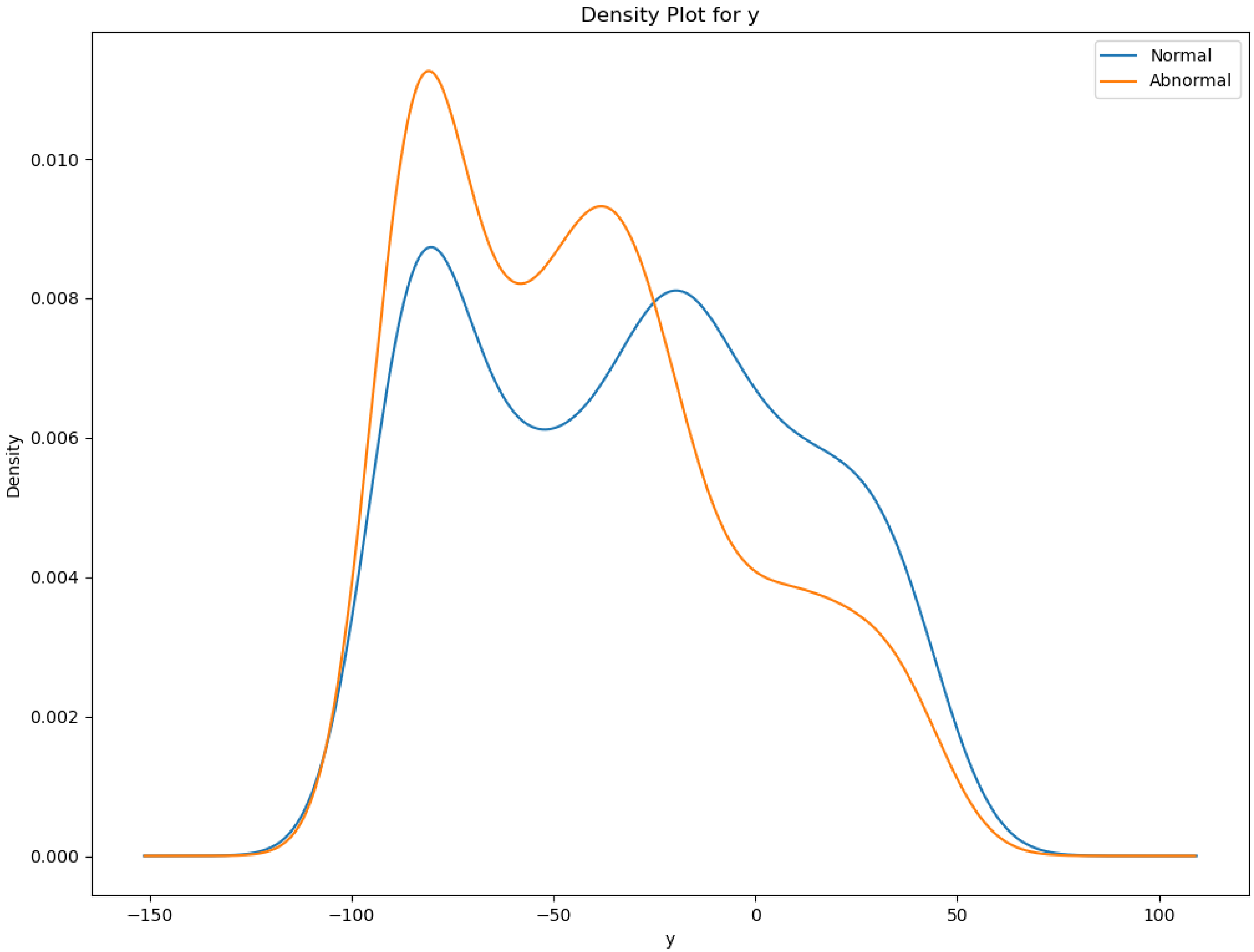
Task 2

Distribution of the Features: Normal vs Abnormal



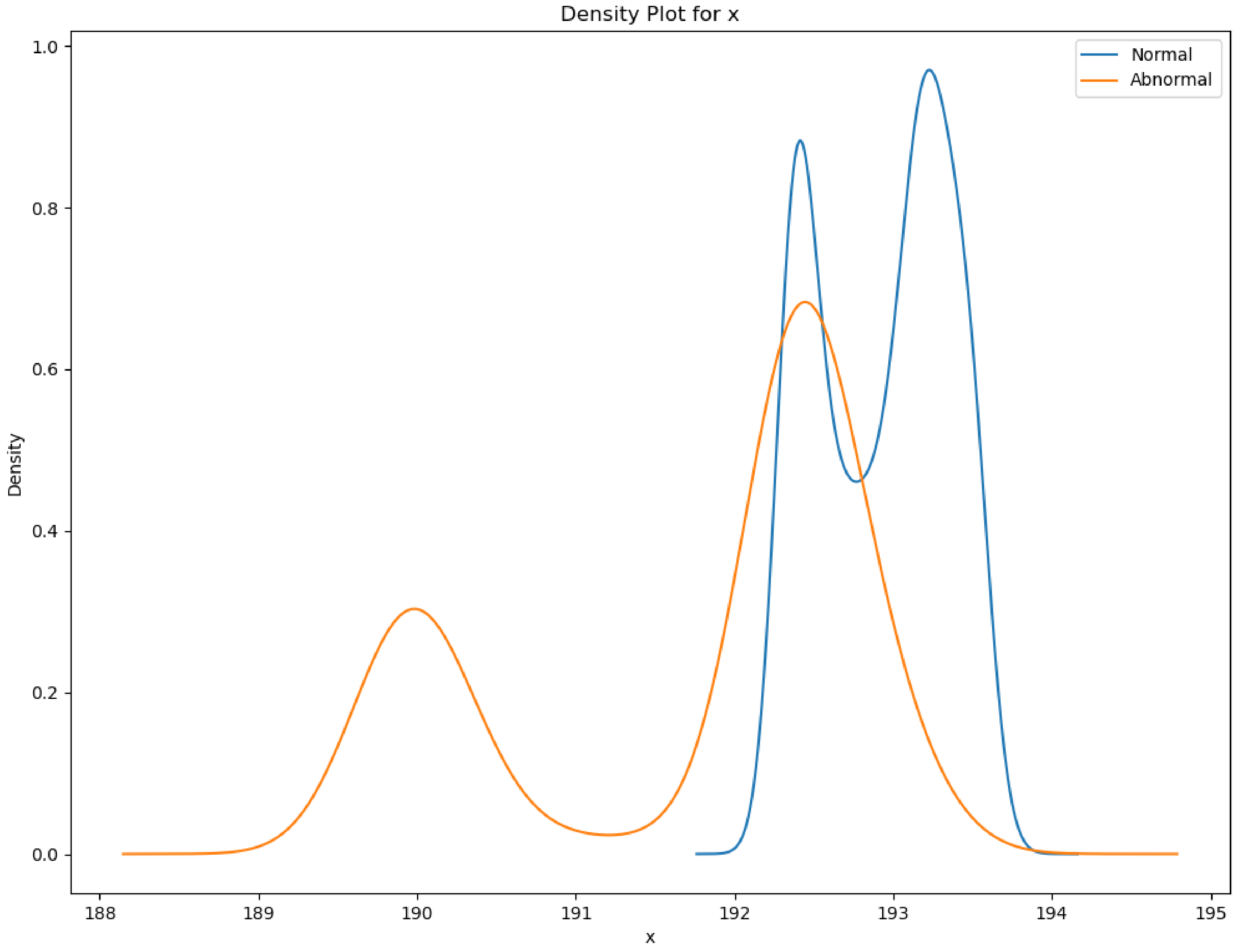
Task 2

Distribution of the Features: Normal vs Abnormal



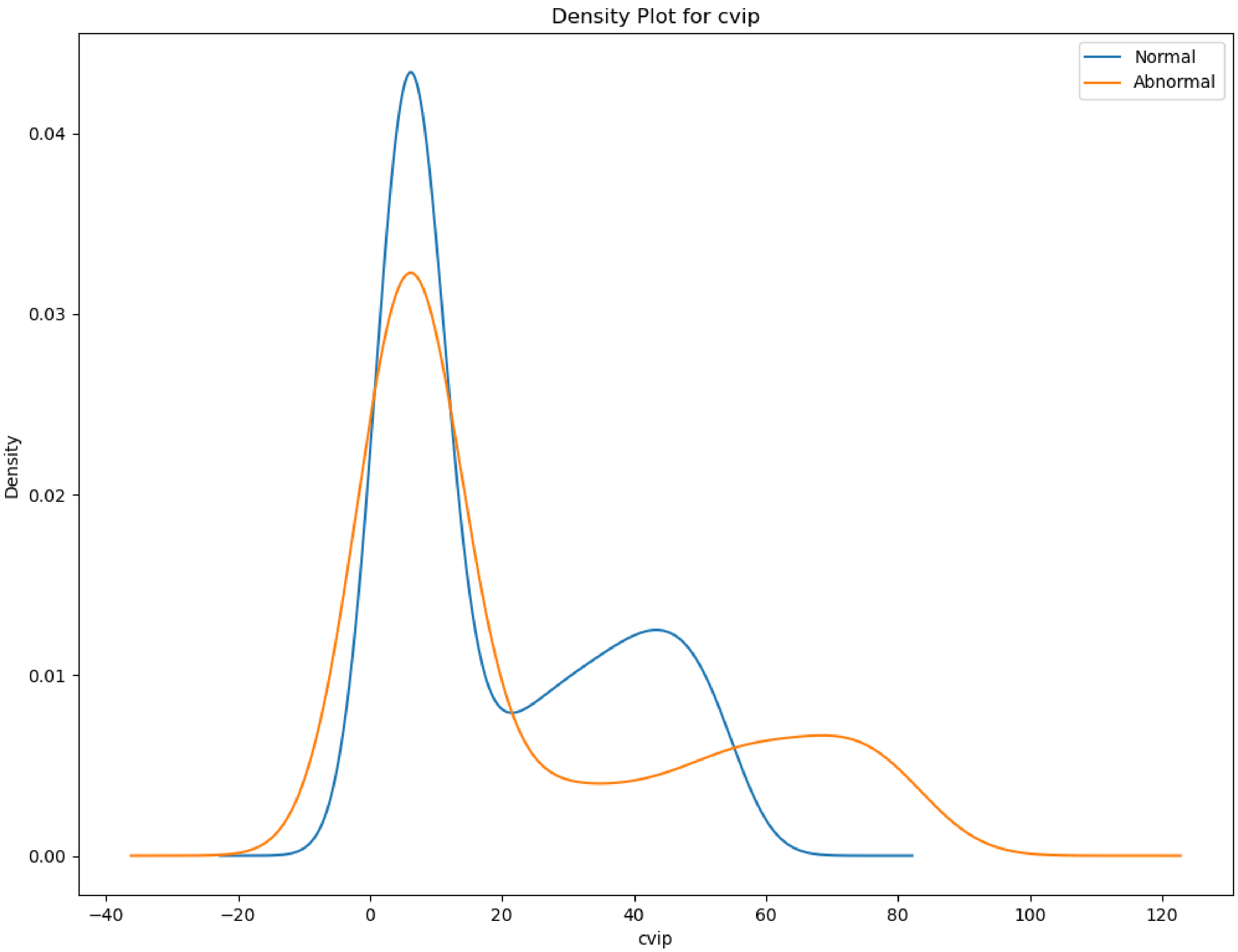
Task 2

Distribution of the Features: Normal vs Abnormal



Task 2

Distribution of the Features: Normal vs Abnormal



Task 2

Q4.b: Use 2-sample t-test to test on the **'steer'** values of **abnormal runs vs normal runs.**

i. State the null and alternative hypotheses

H0: There is no significant difference between the steering variable population means between the normal and abnormal runs.

H1: There is a significant difference between the steering variable population means between the normal and abnormal runs.

ii. Perform the test and calculate test statistics

T-statistic: -2.059574394684549

P-value: 0.03949993360519476

iii. Assume a significance level of 0.05, what is your conclusion

There exists significant evidence to reject the Null Hypothesis.

Q4.c: Does the testing result contradict your observation on the "steer" feature in part 4.a? Why?

This outcome contradicts our intuition as both the normal and abnormal distributions appear to have means centered very close to zero. Thus, we originally thought that there wouldn't be a significant difference between their means.

Task 2

Q5: Some of the features are better indicators of abnormal AV behavior, can you identify them?

a. By looking at the distribution plots of the features in Task 2.4, explain your choice of indicators.

- Steer - The normal runs had a higher density closer to zero and the abnormal had a higher variance and spread.
- x and cvip - The abnormal plot shows two local maximums, one indicating a closer distance to the NPC actor and the other indicating a farther distance which was the response during the clear-night run where the EV changed lanes.

b. For the fields you identified as good accident indicators above, are they related (Calculate the Pearson correlation coefficient between each pair of the indicators to justify your answer)? If so, how does that affect the predicting power of using one indicator versus using all of them?

Pearson Correlations

Steering and x: 0.16049701319159093
Steering and cvip: -0.06898881039039098
cvip and x: -0.3015787487970796

Although we don't have enough collinearity with these features to have a major negative difference in the predicting power, in general we would want to avoid using features with higher correlation since it would take away from the predicting power by skewing the data. Of the above options, cvip and x are the most correlated and it may be redundant to include them both. That being said, since these features are relatively independent, using all of them would give us more predicting power than just using one of them.

Task 2

Q6: Suppose we want to use hypothesis testing to test whether the field you choose from Task2.5 is indeed a good indicator of abnormal AV behavior, using the Kolmogorov–Smirnov two-sample test.

a. Construct the null and the alternative hypothesis and state them below

   H0: There is no significant difference between the x variable distributions between the normal and abnormal runs.
   H1: There is a significant difference between the x variable distributions between the normal and abnormal runs.

b. Perform the KS two-sample test and calculate its statistics.

   T-statistic: 0.5740786464011629
   P-value: 2.6613101312928488e-269

c. Assume a significance level of 0.05, what is your conclusion?
   There exists significant evidence to reject the Null Hypothesis.

d. Repeat the same test on a feature that you did not select as an indicator of abnormal behavior in Task 2.5, what is your conclusion?

- Using ‘brake’ values of abnormal runs vs normal runs
- H0: There is no significant difference between the brake variable distributions between the normal and abnormal runs.
- H1: There is a significant difference between the brake variable distributions between the normal and abnormal runs.

T-statistic: 0.03556296628829175

P-value: 0.210072458272195

Failed to reject Null Hypothesis.

e. What are the major differences between the KS test and the t-test?

Some of the major differences between the t-test and KS test include the following:

- T-test:
  - Measures the difference in means.
  - Used to determine whether two groups have different true means.
  - Statistic calculated with the pooled standard error and the number of observations in each group.
- KS-test:
  - - Measures the maximum difference between CDFs of two samples.
  - - Determines whether samples are drawn from different distributions.
  - - Incorporates number of observations but NOT pooled standard error, or any measure of spread.

The t-test looks squarely at the means, though still taking account for variance (standard error) and group size. The KS-test looks at the distribution as a whole and any large discrepancies within.

Task 2

Q7: Keeping in mind that this experiment is executed over a period of time, what assumption did you make when using the KS two-sample test on the distributions in Task2.6? Are you able to come up with one situation where this assumption fails?

The assumption we took was that all variables are independent to properly take the KS two-sample test on the distributions of task 2.6. This means treating the data as if each moment in the simulation is independent of the previous one when actually it is not. This assumption fails when events happening in the simulation are dependent on the ones that happened before, for example, the braking part of the simulation. As the AV has a reaction time, as he sees the car beginning to cut him off (t-1) in timeframe t he reacts in a way that is dependent on t-1 and initiating breaking, making v, the throttle, braking, and cvip at time t also dependent at time t-1 therefore making this assumption fail.

Task 2

Q8: The dynamic-time-wrapper (DTW) is a method to compare two time-series data (such as the control and the trajectory data collected in our simulation). Use the DTW package in python (dtaidistance · PyPI), and apply the DTW distance on the two time-series dataset (using steering data of clear-noon as a reference): (1) steering data of clear-night and (2) steering data of clear-sunset. What can you say about the DTW distance for (1) and (2) with respect to the reference?

- steering data of clear-night 1.17603264771516
- steering data of clear-sunset 0.050521340023401175

- There is a relatively low DTW distance between the steering data of clear-sunset and the reference compared to clear-night and reference. Thus, the steering data for clear-sunset is more similar to the reference (clear-noon) than clear-night.