Cheryl Leigh, Sophia Akhter, Jacob Chung

Topic Idea #1: Mario Kart Win Prediction

1. **Problem Statement**
   - ✔ We would like to tackle the problem of determining the best Mario Kart settings to choose in order to win my games. Avid Mario Kart fans know that seemingly-minor factors such as vehicle handling or speed under water can make or break a game, and further research can lead to concrete answers about the best characters and vehicles to choose.
   - ✔ Our topic is how to win Mario Kart 8.
   - ✔ We would like to learn more about the best combination of vehicles and characters to be successful in Mario Kart.

2. **Significance of the Problem**
   - ✔ This problem is important because once you become an experienced Mario Kart player and competition is tight, it is important to pay attention to details and find small edges to get ahead and continue to become a better player.
   - ✔ Insights from this project could be helpful because once players know the basics of Mario Kart and are comfortable playing, concrete numerical data could help fill in the gaps in our knowledge about how to win most easily.

   - ✔ https://towardsdatascience.com/the-best-mario-kart-8-character-according-to-data-science-d300dd7a645a

   - ✔ https://medium.com/civis-analytics/the-best-mario-kart-character-according-to-data-science-7dfb65d4c18e

   - ✔ https://lifehacker.com/win-mario-kart-with-the-15-best-character-kart-tire-com-1826778879

   - ✔ https://bicorner.com/2018/07/31/data-science-shows-that-the-best-mario-kart-8-character-is-wario/

3. **Potential Datasets**
   - ✔ Look into some potential datasets using the online resources provided on Canvas or other data sources you find online. This is exploratory, so you won't commit to any specific dataset yet.
   - ✔ You are not expected to identify a dataset yet. This is intended to get you started.
   - ✔ Look through the datasets to identify any potential matches.
   - ✔ Provide links to potential datasets here, if any.
   - ✔ If you cannot find any potential datasets, which is completely fine at this stage, simply explain your process (where you looked, etc.).

   - ✔ https://www.kaggle.com/barelydedicated/mariokart8

   - ✔ https://github.com/woodnathan/MarioKart8-Stats

Topic Idea 2: NYC Traffic Accidents

1. **Problem Statement**

   ✔ The problem that our group would like to tackle is discovering trends in traffic accidents in New York City. Driving is the main source of transportation for many, so being safe and efficient while driving can lead to a higher quality of life.

   ✔ The topic of this project is analyzing where past accidents have occurred and how deadly each one was, and also predicting where future accidents have happened so that people can avoid them. Nobody likes waiting in traffic, and motor vehicle accidents can cause your commute home to be twice as long as it really should be. Understanding where accidents happen and when to avoid them can save oneself a lot of time.

   ✔ We want to learn which months traffic accidents have occured the most in the past, and if there might be ties to a particular season. We also want to know the time and day these accidents occur. For example, rush hour on Friday afternoons could be a point of interest. Also, is there a contributing factor to these crashes such as being on one's phone or running a red light? Is there a particular type of vehicle that gets into crashes the most?

2. **Significance of the Problem**

   ✔ It is important to tackle this problem because we as humans want to be safe. We can prevent accidents by learning about these different conditions and factors that go into a traffic accident.

   ✔ The insights from this project could be useful because people will now know what months, time of the day, day of the week, and street that accidents occur most on. This could also help the city of New York to know where to possibly develop more crosswalks or provide new crosslights to protect pedestrians.

   ✔ https://dmv.ny.gov/about-dmv/archives-statistical-summaries

   ✔ https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2564424/

   ✔ https://www.sciencedirect.com/science/article/abs/pii/0021968162900516

   ✔ https://www.nyc.gov/html/dot/downloads/pdf/nyc_ped_safety_study_action_plan.pdf

3. **Potential Datasets**

   ✔ https://www.kaggle.com/mysarahmadbhat/nyc-traffic-accidents

   This dataset is from kaggle. It has 29 columns which range from crash date, time, borough, street, exact location, people injured/killed, pedestrians injured/killed, cyclists injured/killed, motorists injured/killed, contributing factors, vehicle type, and others.

   ✔ https://www1.nyc.gov/site/nypd/stats/traffic-data/traffic-data-collision.page

   This collection of datasets from the New York City government website has motor vehicle accident data throughout the different boroughs of NYC. There is a citywide data, but also focuses on intersections, highways, bridges, or tunnels.

Topic Idea 3: Cost of Higher Education in US

1. **Problem Statement**
   - ✔ The problem we want to tackle for this project is the high costs of higher education in America while college degrees are needed for a lot of opportunities .

   - ✔ The topic of the project is how unreasonable the costs of college tuition is in the United States.

   - ✔ We would like to learn about the average costs, debt, and demand for college degrees

2. **Significance of the Problem**
   - ✔ I believe it's important to tackle this problem because it's relevant to us as students along with the whole country as a system. Since college tuition is so expensive and without financial aid, student loans, or scholarships, are unreasonably priced.

   - ✔ I think some insights on this project would be the true value of a college education, rethinking the system in which we need a degree to succeed, along with how other countries offer free education and to rethink ours.

   - ✔ https://educationdata.org/average-cost-of-college

   - ✔ https://www.ed.gov/college

   - ✔ https://edsource.org/2020/tuition-free-college-is-critical-to-our-economy/641232

   - ✔ https://www.pewresearch.org/social-trends/2011/08/17/i-the-value-usefulness-and-cost-of-college/

3. **Potential Datasets**
   - ✔ I looked at some websites regarding costs of college tuition

   - ✔ https://www.kaggle.com/jessemostipak/college-tuition-diversity-and-pay

   - ✔ https://data.ed.gov/dataset/college-scorecard-all-data-files-through-6-2020/resources

   - ✔ https://www.kaggle.com/kaggle/college-scorecard/code