Initial Problem

Universal truths:

- F1 races have a combination of factors that can determine success
- Marginal gains across many different factors can add up to make a significant difference
- It is difficult for a human to compile information across multiple sources to make quick decisions
- F1 races require decisions to be made in very short time frames (some 10-30s, others 0-10s)

Problem:

- Mercedes-AMG need to give themselves the strongest edge possible to win games
- They have a system already in place to aggregate data and make decisions but need to constantly improve
- There is currently a gap in that they do not have open radio communications integrated, which means a source of information is not being utilised and acted upon

Knowledge:

- Large Language Models can help to read qualitative data and text at high speed, computationally process, and then give outputs in comprehensible text

Hypothesis:

- An LLM could be utilised to ‘listen’ to radio communications and turn that into actionable information for the Mercedes-AMG F1 team

Research output:

- Any information from the radio communications needs to be easy for humans to digest
    - Think concise, clear, actionable
    - Source: Chatbot
- The AI/LLM used to analyse radio communications needs to be auditable - the team need to be able to understand how it's reached conclusions so it can trust them. It can't be a *Black-box*
    - Instant output: Information and action
        - For example 'ALERT: X has happened/been interpreted. ADVICE: Take Y Action'
    - Further output: Log that explains how information was generated
    - Source: Chatbot
- Reliability is the most important factor for Mercedes team
    - Ensure the AI provides consistent and accurate information. It should handle the variability in data quality, especially during different race conditions.
    - Source: Chatbot
- Radio comms between team will often be encoded, for brevity and to prevent other teams understanding
    - Source: F1 fan, chatbot, Reddit
- It is difficult to know exactly what the codes mean, but there are some telltale signs - a key one being 'Plan X + 7 - the PLUS 7 means it will be executed in 7 laps
    - Source: F1 fan
- Sometimes teams might purposefully create false messages to confuse competitors - '*there was a race last year where Vettel, over the radio, underplayed his tire wear so that the drivers/teams behind him would think they couldn't catch/pass him. When the race was over he admitted, "those tires were DONE!"*'
    - Source: Reddit
- Most messages will not have actionable info - 16% of a random sample contained what appeared to be strategic information
    - Source: OpenF1 analysis
- Certain keywords implied decision-making / changes in strategy:
    - Plan
    - Thinking
    - Do not use
    - Plus
        - Source: OpenF1 analysis
- Many messages had accented voices - 52%, this could make transcription and comprehension by LLMs more difficult
- Many messages had background factors which made them difficult to understand (background noise, radio interference) - 64% had high or medium difficult to hear

Conclusions

- Large language model used to interpret F1 radio comms data, will need to filter through lots of information and need to identify useful information and relay back in an easy-to-digest manner
- **Features**:
    - Need to feed in radio communications
        - Either via transcription to text, or being able to interpret audio files
        - Will need multiple plug-ins to pick up all radio channels
    - Needs to provide outputs in a concise manner - format of:
      ALERT: (Information)
      ADVICE: (Action)
    - Needs to produce a reviewable/auditable log of data it processed, how it computed it and translated to outputs
    - Needs to be able to learn from data on what radio communications mean and what actions to take
    - Need a way for users to be able to manually configure how it is interpreting data to further optimise
- **Implementation issues**:
    - Will need significant amounts of existing data run through, tested and iterated to get it familiar with style of comms and what things may mean
    - May need to train it on how to actually decipher messages given by code in teams, potentially even how to learn different codes for each team
    - There may be difficulty in interpreting accented English - at transcription and AI load-in phase
    - There will likely be difficulty ‘hearing’ messages, due to background noise and radio interference
    - LLMs are prone to ‘hallucinations’ and giving false data - this ties in with a retrospective log that can be used to capture and for operators to configure and prevent in the future