**Analysis Description**

My analysis has come to several conclusions regarding what would make a popular podcast. Based on data from the most popular podcast the most common duration is about an hour with a fair amount of variation depending on the podcast, therefore the ideal podcast length is probably between 45 minutes and 1.5 hours. Moreover, over 50% of the most popular podcasts had over 100 episodes which seems to imply that being consistent and continuing to release new episodes is vital for a new podcast to thrive. As a podcaster it is also important to know your audience and based on Spotify user behavior data most people prefer shorter podcasts to longer ones with about a third of people liking both. Upon further analysis it appears that the frequency that someone listens to a podcast has an effect on whether they want a shorter or longer podcast. People who rarely listen to podcasts want shorter podcasts, people who listen once a week don’t seem to have a preference as do people who listen daily, but people who listen to podcasts several times a week seem to have a preference for longer podcasts. New podcasters should understand who their audience is so they can tailor the length of their podcast appropriately. It is also important for new podcasters to understand whether their audience is majority men or women as this has an effect on the preference of podcast format, genre, and host preference. Among the formats of Storytelling, Conversational, Interviews and Educational men seem to have nearly even split among all 4 formats while women have a strong preference for Storytelling over Interviews and Educational. When it comes to Host Preference it seems that men have a sizeable preference for Well Known podcasters over unknown whereas women are split 50-50 between those two categories. Finally among the genres (many were cut out due to lack of reported data for those genres) of Sports, Lifestyle and Health, Health and Fitness, Food and Cooking, and Comedy men and women are pretty even on most categories except that men seem to prefer the genre Sports over Health and Fitness and Lifestyle and Health while women seem to prefer the reverse. All this should provide new podcasts with ideas of how to tailor their content to their audiences and all this will be expanded upon in the next steps of this process.

**Analysis Documentation**

* Spotify Top 100 USA Podcasts with EP
  + top\_pocast\_episode\_ranks
    - Calculate podcast duration (all in minutes)
      * Mean: 71.69354577
      * Median: 54.07346667
      * Mode: 480.7966333
      * Standard Deviation: 82.40537947
    - Calculate number of explicitly true = 10592
    - Calculate number of explicitly false = 17661
  + top\_pocast\_ranks
    - Calculate the number of podcast episodes each podcast has in the top\_pocast\_episode\_ranks sheet using the COUNTIF functions for all top 100 podcasts
      * Found out that this number is already calculated in the total episodes column 😅
    - Calculate number of “mixed” media type podcasts = 12
    - Calculate number of “audio” media type podcasts = 88
    - Calculate number of top 100 podcast with 100 episodes or more = 51
    - Calculate number of top 100 podcast with 10 episodes or more = 79
    - Calculate number of top 100 podcast with less than 10 episodes = 21
* Spotify User Behavior Dataset
  + Calculate preffered\_podcast\_duration numbers
    - Both: 117
    - Longer 81
    - Shorter: 172
    - None: 20 (REMOVED IN PIVOT TABLE)
  + Calculate pod\_host\_preference
    - Both: 171
    - Well Known individuals: 104
    - Unknown Podcasters: 84
    - None: 31 (REMOVED IN PIVOT TABLE)
  + Calculate preffered\_pod\_format
    - Story Telling: 144
    - Converstational: 99
    - Interview: 73
    - Educational: 46
    - None: 28 (REMOVED IN PIVOT TABLE)
  + Calculate fav\_pod\_genre
    - Comedy: 104
    - Lifestyle and Health: 101
    - Health and Fitness: 77
    - Sports: 50
    - Food and Cooking: 15
    - None: 29 (REMOVED IN PIVOT TABLE)
    - Other: 14 (REMOVED IN PIVOT TABLE)
  + 4 pivot tables were made with gender as columns for all 4 and the rows of each table being: pod\_host\_preference, fav\_pod\_genre, preffered\_podcast\_duration, and preffered\_pod\_format
    - I changed the values to be shown as % of column as this allows us to see insights in what % of each gender has to show
    - The column of ‘Other’ was removed as it was small compared to ‘Male’ and ‘Female’ and it makes our analysis cleaner
  + 4 more pivot tables were made with pod\_lis\_frequency as columns for all 4 and the rows of each table being: pod\_host\_preference, fav\_pod\_genre, reffered\_podcast\_duration, and preffered\_pod\_format
    - I changed the values to be shown as % of column as this allows us to see insights in what % of each listener frequency each pivot table has to show
  + None values were removed across Pivot tables mentioned above as while it is data it does not provide insight relevant to the analysis