Spotify Playlists

Share and discuss your responses to each of the following questions with your group.

1. Do you have Spotify or some other digital music player? Have you used the shuffle play feature? If you have used the shuffle play feature, have you ever wondered how truly random it is?

2. What comes to mind when you hear the word, ‘random’?

3. If Spotify is not producing a random sequence of songs, then what might the sequence of songs look like? What would you expect to see?

4. Do you think you can be 100% certain that a sequence of songs was not randomly generated? Explain your answer.

Group Task

Albert Hoffman, a Spotify user, has tweeted to @spotifycares to complain about the shuffle play feature. He writes that every day he takes an hour-long walk and listens to Spotify using the shuffle play feature. **He believes that the shuffle play feature is producing playlists in which some artists are played too often and others are not played enough.**

He has claimed that the Spotify shuffle play feature is not generating random playlists. As evidence, Mr. Hoffman has provided both his music library (8 artists with 10 songs each) and three playlists (20 songs each) that Spotify generated using the shuffle play feature.

Daniel Ek, the CEO of Spotify, has contacted your group to respond to Mr. Hoffman's complaint. He has provided your group with several playlists of 20 songs each using the same songs as Mr. Hoffman's library but generating them using a genuine random number generation method.

To help your group respond to Mr. Hoffman, the next four sections of the problem are designed to help your group explore properties of the randomly generated lists to develop rules that could help determine whether a set of playlists provide evidence that the shuffle feature is not producing randomly selected songs.

Explore and Describe

Examine the 25 randomly generated playlists (available at https://zief0002.github.io/statistical-thinking/spotify/spotify-training-playlists.html) to get an idea of the characteristics of these lists. Write down at least two characteristics about the randomly generated playlist that help you address Mr. Hoffman's concern.

Develop Rules

Use the set of characteristics that your group wrote down to describe randomly generated playlists in the previous section to create a set of one or more rules that flag playlists that **do not appear to have been randomly generated.** (Be sure that each of the characteristics in the previous section is included in a rule.) *These rules should be clearly stated so that another person could easily use them.*

Test Rules

Use five additional randomly generated playlists to test your rules (available at https://zief0002.github.io/statistical-thinking/spotify/spotify-validation-playlists.html). Let your instructor know that you are ready to receive these playlists. See whether the set of rules your group generated would lead someone to (incorrectly) question whether these playlists are not randomly generated. Based on the performance of your group's set of rules, adapt or change the rules as your group feels necessary.

Evaluate

Use Mr. Hoffman's original three playlists (available at https://zief0002.github.io/statistical-thinking/spotify/hoffman-playlists.html) to apply your group's rules in order to judge whether there is convincing evidence that Mr. Hoffman's Spotify shuffle play feature is producing playlists which do **not** seem to be randomly generated.

Summarize

Your group will now write a letter to Mr. Hoffman that includes the following:

- Your group's set of rules, used to judge whether a playlist does not appear to have been randomly generated. In your letter the rules need to be clearly stated so that another person could apply them to a playlist of 20 songs from Mr. Hoffman's music library;
- A response to Mr. Hoffman's claim that the shuffle play feature is **not random** *because it produces playlists in which some artists are played too often and others are not played enough.*

Type the letter in a word-processed document and email it to each of your group members and the instructor.

Discussion

As a group, discuss your responses to each of the following questions.

5. What made it difficult to come up with a rule to determine whether a sequence of data had been randomly generated? Explain.

6. How might your rules *change* if there were **not an equal number of songs for each artist**? Or a longer set of songs per playlist? (Be specific about how your rules might change.)

7. What does your group need to do to improve the process of working as a team? Be specific about how each member of the group will contribute to this improvement.