MESSAGE TO DANCERS

Hello everyone,

Thank you very much to everyone that sent his or her step data. This work helped me learn to do the following:

1) Excel – Clean and Manipulate Data
2) SQL – Clean, Manipulate, Model and Reporting
3) Power BI – Create Visualizations/Insights

BACKGROUND & QUESTIONS

While on a conversation with my girlfriend, Alex, she was distraught that she did not hit her step goal for the day, which led me to look over my step count. I noticed immediately the only steps per day that was over 10,000 were weekends and those weekends were "Music festival" days. Music Festivals are a big music event where artists, ranging from 50 to 200+, are gathered to perform their latest music/playlist to fans across the world. As I further explore, certain times had significantly more steps than the others, which led me to ask two questions:

1) At music festivals, can step data determine which artist(s)/set is more favorable/enjoyable? Least favorable?
2) At music festivals, are individuals more likely to engage and discover artists if they were informed of their steps they devoted to those artist(s)/set?

DATA & VISUALIZATION

Step data was manipulated and visualized in the following way:

1) **Dates** - Friday, Sept 2nd – Sunday, Sept 4th 2022
2) **Location** - Electric Zoo Festival 2022, Randall's Island, NYC (EZOO)
3) **Comparison of Steps, charts:**
    a. Total Day Steps vs. EZOO Steps (Actual)
    b. Total Day Steps vs. Dancer
    c. EZOO Steps / Artist
    d. EZOO Steps / Day
    e. EZOO Steps / Gender

4) **Card Highlights** (*Interactive***)*
    a. Actual "EZOO Steps" Actual (*dancer selectable*)
    b. Average Steps (*dancer, day selectable*)
    c. Top/Artist Set (*dancer, day, Artist selectable*)
    d. Top Day (*dancer, day, Artist selectable*)
    e. Artists Seen (*dancer, day selectable*)
    f. Est. Sum of Duration (*dancer, day, Artist selectable*)

***This Dashboard is interactive, but due to limited basic version of Power BI, PDFs were only made possible*

Upon creating this, there were some constraints when modeling the data:

1) **Location** – this is a huge constraint because knowing exactly where each dancer was when step count was recorded is crucial in knowing which artist was actually stepped at. So the following was done:
    a. *Set attended* – Since the dancers told me which set they went to, the model is assuming they attended majority of the artist(s)/set. If two or more sets were attended around the same time, step count was divide into the sets evenly
    b. *Sum of Duration* – This is estimated because location is not exact

2) **Device & Position** – which device? Where is the device location?
    a. Dancers were not asked which device they used (phone/ watch**)
    b. Position of device was not asked – pants pocket/floor/bag etc.

3) **Festival** - Weathering and any incidents that happened at EZOO is not considered in this dataset (sunny vs. rainy / dancing vs. sitting vs. medical attention etc.)

***Article on whether one device performs better than the other is invalid and data is relative enough to compare which ever is recorded*

INSIGHTS

1) **Favorite Artist(s)/Set** – The data suggests Carl Cox was a top favorite amongst all dancers.
    a. Of the total "EZOO Steps", which is about 197K amongst all dancers, Carl Cox accounts for about 13% steps, which is more than double any other artist/set seen by all dancers
    b. Average of about 5,000 steps, amongst all dancers, were accounted for Carl Cox

2) **Favorite Day** – This insight is tricky. Lets take a look:
    a. "EZOO Steps" chart suggests the top day is Saturday because the total sum of steps from all dancers on Saturday > Friday & Sunday
    b. However, we must refer to the question "can step data determine which artist(s)/set is more favorable/enjoyable?"
    c. Answer – Collectively, Carl Cox received the most steps and his set was on Sunday. Therefore, the card highlighted "Top Day" is noted as Sunday instead

3) **Honorable Mention** – Artists to consider (By Artists/Set) Chart
    a. Dancers, collectively, stepped >10,000 steps for the following artist:
        i. John Summit, Diesel, Gordo, Bijou, Camelphat, Loco Dice, Chris Lake, Will Clarke
    b. While these artists did not claim the top spot (Carl Cox), the data suggests these artists are just as favorable with the amount of steps devoted to each
    c. Based on step data, these artists are more likely to be seen again than their remaining counterparts below them

4) **OTHER** – insights that don't necessarily answer the question
    a. **Artists seen** – 29 distinct artists seen
        i. Collectively, all dancers saw about 29 different artist(s)/set out of 101 (28.7%)
        ii. In dollars, each ticket for the event ranged from $250-$350 (fees and tax included). On average, all dancers spent about $10.34 to see each individual artist

    b. Big Steppa
        i. Amongst all dancers, throughout the weekend (inside and outside the festival), Jackie Pertillo stepped +20% more than all other dancers

Conclusion

Data suggests Carl Cox is the favorable artist/set based on the steps recorded during his set. Sunday is the best day even though more steps were produced on Saturday. To further validate the data, user feedback from each dancer's experience may be needed to conclude favorability.