Tour Data Analysis and Crowd Energy Prediction

Venue Specific Findings

Venue Alpha

- This is a converted monastery. There is a volume limit here where the energy dies down after a certain volume level (around 6).
- Changing the outfit does not seem to have an effect here with each outfit resulting in similar mean crowd energies.

Venue Beta

- This is a gothic nightclub. This ‘comes alive’ at night, with late night shows having higher crowd energies.
- This also has significantly better weekend shows compared to other venues.
- Spandex performs better than other outfits here.

Venue Gamma

- This is a snob pit. This performs better than other venues in all weather conditions.
- Expensive and free shows have higher energy here. This venue is price sensitive.
- Spandex performs better than other outfits here.

Venue Delta

- This is a mosh pit. Higher volume levels lead to more energy here.
- Higher crowd size leads to higher crowd energy. It feeds itself.
- It has higher energy in cheaper shows. This is also price sensitive as the energy goes down with price.
- All outfits do better than at other venues, with spandex and leather doing better than denim.

Verification of the Lead Singer's Scribbles

Hypothesis	**True/False**	**Reason**
Tuesday shows are 'cursed'.	False	There didn't seem to be a correlation between day of the week and energy, with the mean of each day being almost the same
Venue beta (gothic nightclub) comes alive at night.	True	We saw higher crowd energies during late night shows.
Weekend shows are better	True	Crowd energies during weekends were almost always higher for all venues.
Best shows during full moon	False	Crowd energies for every phase of the moon were similar
Rain leads to bad shows	False	For the same venue, weather conditions didn't seem to have much effect.
Venue delta suffers in storms because it is outdoors.	Maybe	Venue delta performs as well as alpha and beta, with gamma performing well in all weather conditions. So either all three of them are outdoors or the reason is something else.

Mosh pits are legendary in the rain.	False	Venue delta didn’t seem to show any improvement over other venues in the rain.
Louder is better at mosh pits	True	Increasing volume showed higher crowd energies.
Mosh pit feeds itself	True	Higher crowd size leads to more energy
Monks have a noise limit	True	Venue alpha (monastery) showed a decrease in energy with volume level.
Some venues are price sensitive	True	Venue gamma and delta are price sensitive. Gamma responds well to higher prices. Delta is the opposite.
Sweet spot between energy and price	True	The price range of 40-80 seemed like the sweet spot for most venues.
Outfits affect crowd energy	Maybe	Changing outfits didn't seem to create a very big effect for a given venue.

Model Choice and Justification

I chose the **Random Forest Regressor** to predict crowd energies.
Why?

- It handles non linearity. Models like linear regression would not be able to handle the non linear relationships like ticket price vs crowd size, etc.
- It learns by making several decision trees and branches. So it is able to learn complex interactions like “venue is delta, price is low, so high energy”.
- It handles noise much better than linear regression by putting them into separate branches so they do not interfere with the general predictions.