Q2) Codes as a python file and Results (contained in the Q2 pitches test probabilities csv) attached

a) This would have been more accurate if I could use the random effects of Pitchers, Catchers and Batters. But due to the amount of unique sets, my laptop was not able to handle it using any packaged mixed model libraries that I tried in Python, Hence the Xgboost. My next try to get it to be more accurate would have been using the Bayesian softmax multilevel modeling in Pymc3, so that we could gain some uncertainty intervals as well as Dirichlet distribution to signify prior belief of these pitches often being a Fastball. For this analysis, I continue to use Xgboost to utilize different factors as a boosted tree to get the values. This gives out the probability of each of those pitches which is in the file I am attaching. This Xgboost has been tuned using the RandomizedSearchCV.

b) Regarding other information that could help me arrive at a more accurate prediction is the sequence of pitches, or at least what was the previous pitch. We can then use game situation as well as the presence of the understanding of sequences to predict what will be the current pitch. The reason behind it is that game situation is not the only thing that effects what pitch will be thrown, because if the batter has already seen two fastballs, regardless of the game situation, throwing another one might be not as fruitful. This assumption can be made with the current dataset, though having a game id or atbat id to identify the ones that are definitely from a single at bat would be more helpful.

Q**3) Codes attached as a R file, here is the ranking.**

Used Mixed effect models to vary intercepts for each VendorID when trying to predict the HotDogSales. This gives me the effect of each vendor towards the hotdogsales.

VendorID	VendorEffectToSales
20	29.1648117
6	17.6538705
3	14.2175339
21	13.7420671
29	11.3723238
16	10.0204979
24	9.9780042
27	7.9162084

5	6.6619208
2	5.8325309
15	5.1698586
28	3.0445034
14	1.5842372
30	1.1529834
9	-0.9884758
10	-2.0143877
11	-2.5949240
17	-3.3278453
22	-3.7090714
25	-5.1419899
19	-5.1677553
23	-7.0933070
7	-8.4947826
8	-8.8163956
18	-10.2872231
13	-12.2929374
1	-13.5063267
26	-13.6124434
4	-19.2439891
12	-21.2194977