Please read this document **very carefully** as it explains the format of the test prediction files you would submit for your projects. The accuracy evaluations are done through a computer program and if your submitted files do not follow these guidelines, it can easily cause a problem and result in your project being excluded from the competition.

File Name

Each team works on two projects. After you did the predictions, a member from the team would email me (aaghasi@gsu.edu) the prediction files. The submitted file **must be in csv format**. Avoid sending other formats such as xlsx, dat, etc. Also depending on the project you pick, the prediction file can **only** have the following names:

Gotham Cabs → G.csv
Amusement Park → A.csv
Justice League Housing → J.csv
Baseball Faces → B.csv

For example, if you have worked on the Gotham Cabs project and you are submitting the prediction file, the file name **must be** G.csv.

How to Submit?

A member from the team would email me the two prediction files. The subject of your email should be like this:
"Prediction Files for Team [your team number]"

To find your team number, please look it up on the schedule link:
https://docs.google.com/document/d/1xy7_oE67jqHnX_sDzuIGmg7Hg8eiSVY5Outpteqw2-Q/edit
The team number is simply the index on the first column.

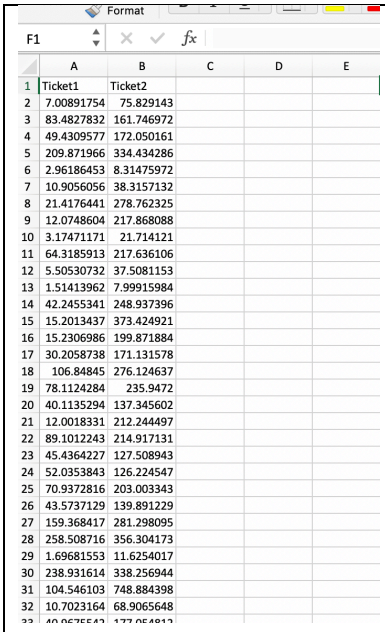
For example, if the "Justice League" team composed of Bruce Wayne, Diana Prince and Clark Kent are going to submit their prediction files, one of them (say Bruce Wayne) would send me an email with the following subject:
"Prediction Files for Team 0"

Note that only one submission is accepted, if you send me multiple emails, they would not be accepted. Please revise your submission before you send it to me to avoid such complications.

File Format

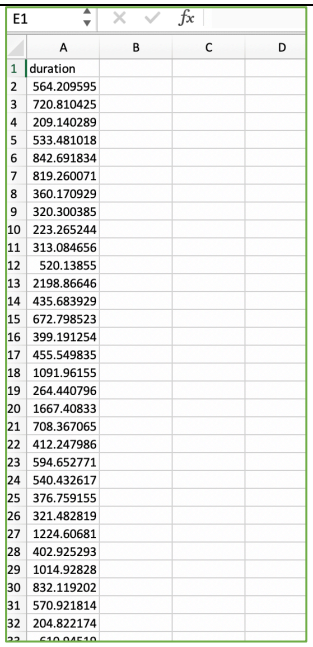
The csv file that you submit should only contain the response variable(s). **The order of the rows should be exactly the same as the order of the rows in the test file.** That basically means that the first row (after the target name) corresponds to the predicted value for the first row of the test file, the second row corresponds to the second, and so on.
As noted above, the prediction files should only contain the response variables. The first row of the csv file would be the name of the response variable and the remaining rows would be the

predicted values. **Failure to follow this format would cause a failure in compiling your file**. To avoid confusions, I have taken screenshots of the first few rows of the files you would need to send me (your files should look similar, only the predicted values should be different). Make sure the response names and orders are **exactly** as shown in these figures.



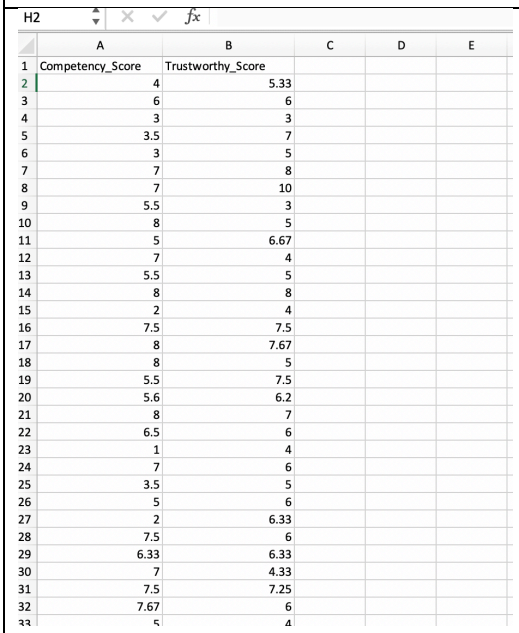
	A	B	C	D	E
1	Ticket1	Ticket2			
2	7.00891754	75.829143			
3	83.4827832	161.746972			
4	49.4309577	172.050161			
5	209.871966	334.434286			
6	2.96186453	8.31475972			
7	10.9056056	38.3157132			
8	21.4176441	278.762325			
9	12.0748604	217.868088			
10	3.17471171	21.714121			
11	64.3185913	217.636106			
12	5.50530732	37.5081153			
13	1.51413962	7.99915984			
14	42.2455341	248.937396			
15	15.2013437	373.424921			
16	15.2306986	199.871884			
17	30.2058738	171.131578			
18	106.84845	276.124637			
19	78.1124284	235.9472			
20	40.1135294	137.345602			
21	12.0018331	212.244497			
22	89.1012243	214.917131			
23	45.4364227	127.508943			
24	52.0353843	126.224547			
25	70.9372816	203.003343			
26	43.5737129	139.891229			
27	159.368417	281.298095			
28	258.508716	356.304173			
29	1.69681553	11.6254017			
30	238.931614	338.256944			
31	104.546103	748.884398			
32	10.7023164	68.9065648			

Figure 1. A partial screenshot of the file A.csv



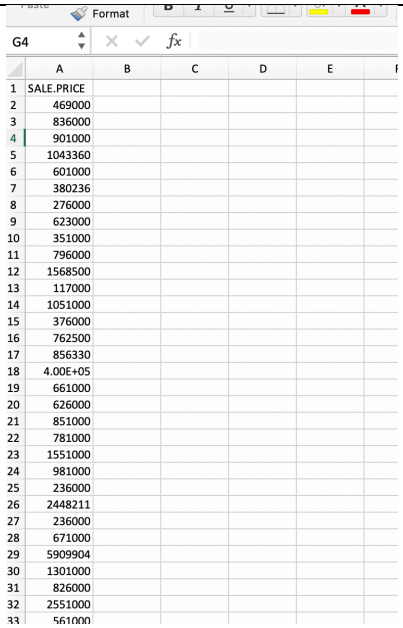
	A	B	C	D
1	duration			
2	564.209595			
3	720.810425			
4	209.140289			
5	533.481018			
6	842.691834			
7	819.260071			
8	360.170929			
9	320.300385			
10	223.265244			
11	313.084656			
12	520.13855			
13	2198.86646			
14	435.683929			
15	672.798523			
16	399.191254			
17	455.549835			
18	1091.96155			
19	264.440796			
20	1667.40833			
21	708.367065			
22	412.247986			
23	594.652771			
24	540.432617			
25	376.759155			
26	321.482819			
27	1224.60681			
28	402.925293			
29	1014.92828			
30	832.119202			
31	570.921814			
32	204.822174			

Figure 2. A partial screenshot of the file G.csv



	A	B	C	D	E
1	Competency_Score	Trustworthy_Score			
2	4	5.33			
3	6	6			
4	3	3			
5	3.5	7			
6	3	5			
7	7	8			
8	7	10			
9	5.5	3			
10	8	5			
11	5	6.67			
12	7	4			
13	5.5	5			
14	8	8			
15	2	4			
16	7.5	7.5			
17	8	7.67			
18	8	5			
19	5.5	7.5			
20	5.6	6.2			
21	8	7			
22	6.5	6			
23	1	4			
24	7	6			
25	3.5	5			
26	5	6			
27	2	6.33			
28	7.5	6			
29	6.33	6.33			
30	7	4.33			
31	7.5	7.25			
32	7.67	6			

Figure 3. A partial screenshot of the file B.csv



	A	B	C	D	E
1	SALE.PRICE				
2	469000				
3	836000				
4	901000				
5	1043360				
6	601000				
7	380236				
8	276000				
9	623000				
10	351000				
11	796000				
12	1568500				
13	117000				
14	1051000				
15	376000				
16	762500				
17	856330				
18	4.00E+05				
19	661000				
20	626000				
21	851000				
22	781000				
23	1551000				
24	981000				
25	236000				
26	2448211				
27	236000				
28	671000				
29	5909904				
30	1301000				
31	826000				
32	2551000				

Figure 4. A partial screenshot of the file J.csv

Submission Deadline

The deadline to submit your prediction files is **Wednesday May 1st, at 10:00 AM Eastern time**. **No files would be accepted after this deadline.** Please try to submit early to avoid any last-minute email delivery issues.