Specifications of generated tables and attributes

Strategy to battery discharge

Step1

We choose the top n peaks (n days) in each season, and choose the hours with the largest k forecasted probabilities on each of these days.

Step 2

We adjust the chosen probabilities to have them sum up to one for each day.

Step 3

We discharge the battery in proportional to the adjusted probabilities. Since the capacity of the battery is $2-$hours, the maximum amount that can be discharged is 0.5. If the probability is $p\%$, we discharge $p\%$ of 0.5 capacity. Then the discharged energy is $p\%$ of 0.5.

If one of the probabilitis on some day is larger than 0.5 , i.e., $p > 0.5$, we only discharge 50% of 0.5 capacity, and spread out the remaining capacity in the rest hours evenly.

For example, if we use the 3 largest probabilities on each day, and the probabilities are $0.6, 0.1, 0.3$, then we discharge $0.5/0.5 = 100\%$, $0.25/0.5 = 50\%$, $0.25/0.5 = 50\%$ at each hour, repectively.

Step 4

We eavaluate the performance of this strategy by using the ratio of the total energy we successfully discharge and the total number of peaks we use. The more closer this ratio is to 100%, the better performance we have.

Table "season_results"

season_results

	adjusted_demand_MW	demand_MW	season	ts	rankings_per_day	rankings_per_season	forecast
5	20702.0	21168.0	2017-2018	2017-06-12 17:00:00	1.0	3.0	0.3025
6	20122.0	20536.0	2017-2018	2017-07-19 18:00:00	1.0	5.0	0.2125
7	21170.0	21786.0	2017-2018	2017-09-25 17:00:00	1.0	1.0	0.66
8	21039.0	21542.0	2017-2018	2017-09-26 17:00:00	1.0	2.0	0.72
9	20238.0	20906.0	2017-2018	2018-01-05 18:00:00	1.0	4.0	0.4675

Table "season_results" shows the ground-truth **top 5 peaks** in the **season 2017-2018**, with **3 peaks at 17:00:00** and **2 peaks at 18:00:00**.

- **demand_MW**: demand in Megawatts for each hour
- **adjusted_demand_MW**: demand after adjusting in Megawatts for each hour
- **season**: year-based cycle, from last year's Mar. 31 to next year's Apr. 1
- **ts**: date-time
- **rankings_per_day**: rankings of adjusted demand in a day
- **rankings_per_season**: rankings of adjusted demand in a season
- **forecast**: forecasted probability of being the peak

hour on that day

Table "season_adjusted"

Table "season_adjusted" shows the **3 largest** forecasted probabilities on each day of the **top 5 peaks** in the **season 2017-2018**.

season_adjusted

	ts	forecast	prob_rankings_per_day	is_true_peak	adjusted_prob	discharge_rate	season	top_n
1197	2017-06-12 17:00:00	0.3025	2.0	1	0.36	0.72	2017-2018	5
1198	2017-06-12 18:00:00	0.3125	1.0	0	0.38	0.76	2017-2018	5
1199	2017-06-12 19:00:00	0.215	3.0	0	0.26	0.52	2017-2018	5
1310	2017-07-19 17:00:00	0.515	1.0	0	0.5	1.0	2017-2018	5
1311	2017-07-19 18:00:00	0.2125	2.0	1	0.25	0.5	2017-2018	5
1312	2017-07-19 19:00:00	0.1625	3.0	0	0.25	0.5	2017-2018	5
1515	2017-09-25 17:00:00	0.66	1.0	1	0.5	1.0	2017-2018	5
1516	2017-09-25 18:00:00	0.195	2.0	0	0.25	0.5	2017-2018	5
1517	2017-09-25 19:00:00	0.0525	3.0	0	0.25	0.5	2017-2018	5
1518	2017-09-26 17:00:00	0.72	1.0	1	0.5	1.0	2017-2018	5
1519	2017-09-26 18:00:00	0.15	2.0	0	0.25	0.5	2017-2018	5
1520	2017-09-26 19:00:00	0.0625	3.0	0	0.25	0.5	2017-2018	5
1825	2018-01-05 18:00:00	0.4675	1.0	1	0.47	0.94	2017-2018	5
1826	2018-01-05 19:00:00	0.44	2.0	0	0.44	0.88	2017-2018	5
1827	2018-01-05 20:00:00	0.0875	3.0	0	0.09	0.18	2017-2018	5

- **prob_rankings_per_day**: rankings of forecasted probabilities in a day
- **is_true_peak**: if this hour is the ground-truth peak

- **adjusted_prob**: adjusted probabilities in the selected top k probabilities (here, $k = 3$)
- **discharge_rate**: energy discharged in proportional to adjusted probabilities, i.e., probability/0.5
- **top_n_peaks**: number of top peaks to use

  The three peaks at **17:00:00** in the **season 2017-2018** all appear in the table "season_adjusted", so all three are hitted, the hit rate is **3/3** for **"season: 2018-2018, hour: 17, top_n_peaks: 5"**.

  The two peaks at **18:00:00** in the **season 2017-2018** both appear in the table "season_adjusted", so both are hitted, the hit rate is **2/2** for **"season: 2017-2018, hour: 18, top_n_peaks: 5"**.

  The total hit rate for **"season: 2017-2018, top_n_peaks: 5"** is $5/5(100\%)$.

Table "hour_hits"

hour_hits

	adjusted_demand_MW	demand_MW	season_x	ts	rankings_per_day	rankings_per_season	forecast_x	prob_rankings_per_day	is_true_peak	adjusted_prob	discharge_rate	top_n
0	20702.0	21168.0	2017-2018	2017-06-12 17:00:00	1.0	3.0	0.3025	2.0	1	0.36	0.72	5
1	21170.0	21786.0	2017-2018	2017-09-25 17:00:00	1.0	1.0	0.66	1.0	1	0.5	1.0	5
2	21039.0	21542.0	2017-2018	2017-09-26 17:00:00	1.0	2.0	0.72	1.0	1	0.5	1.0	5

Table "hour hits" shows the hitted peaks for **"season: 2017-2018, hour: 17, top_n_peaks: 5"**. The averaged discharged energy is $(0.72 + 1. + 1.)/3 = 91\%$, and is displayed in the bracket under the **column "17"** for **season "2017-2018"** in the table "Report" .

hour_hits

	adjusted_demand_MW	demand_MW	season_x	ts	rankings_per_day	rankings_per_season	forecast_x	prob_rankings_per_day	is_true_peak	adjusted_prob	discharge_rate	top_n
0	20122.0	20536.0	2017-2018	2017-07-19 18:00:00	1.0	5.0	0.2125	2.0	1	0.25	0.5	5
1	20238.0	20906.0	2017-2018	2018-01-05 18:00:00	1.0	4.0	0.4675	1.0	1	0.47	0.94	5

Table "hour hits " shows the hitted peaks for **"season: 2017-2018, hour: 18, top_n_peaks: 5"**. The averaged discharged energy is $(0.5 + 0.94)/2 = 72\%$, and is displayed in the bracket under the **column "18"** for **season "2017-2018"** in the table "Report" .

Report

Report

	Season	Top_n_peaks	Top_n_probs	HitRate(%)	Performance(%)	17	18
0	2017-2018	5	3	5/5(100.0%)	83.0	3/3(91.0%)	2/2(72.0%)

Table "Report" shows the report for**"season: 2017-2018, top_n_peaks_peaks: 5"**.

- **HitRate(%)**: $x/y(p\%)$ means we discharge the battery successfully for x out of y peak hours. $p\% = x/y * 100\%$.
- **17**: $x/y(p\%)$ means we discharge the battery successfully for x out of y peak hours which appears at 17:00:00 and the average discharged energy for each of these x peak hours is $p\%$.
- **18**: $x/y(p\%)$ means we discharge the battery successfully for x out of y peak hours which appears at 18:00:00 and the average discharged energy for each of these x peak hours is $p\%$.
- **Performance(%)**: displays the ratio of the total energy we successfully discharge and the total number of peaks we use, i.e., $(91\% * 3 + 72\% * 2)/5 = 83.4\%$. Since the number of peaks we use (i.e., 5) is fixed, to make the performance 100%, we need to successfully

discharge 100% for each of these 5 peaks. Therefore, the more closer to 100%, the better performance.