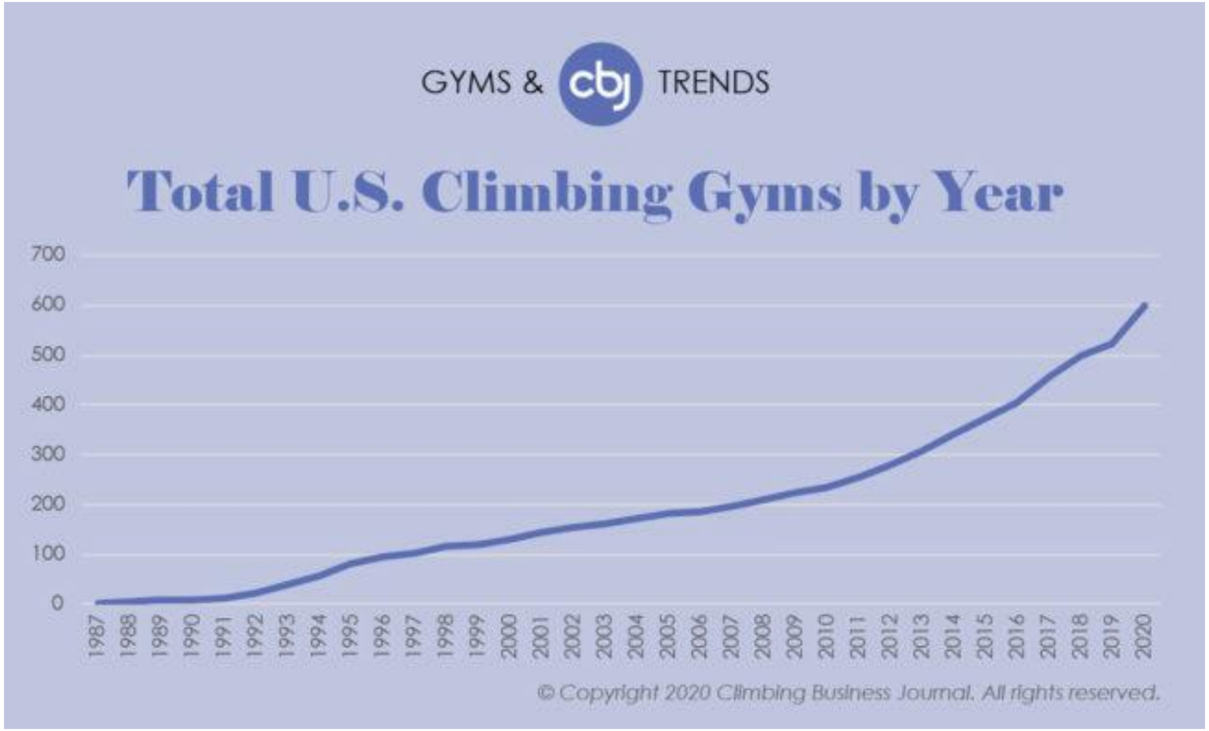
Olympic Climbing Prediction

Hanna Landrus

Climbing Rising Popularity



https://www.climbingbusinessjournal.com/climbing-gyms-and-trends-2019/

Competitive Climbing History

1989 - First World Cup climbing competitions series held

2007 - International Federation of Sport Climbing (IFSC) founded and began hosting world cups

2016 - International Olympic Committee announced climbing added to 2020 Olympics

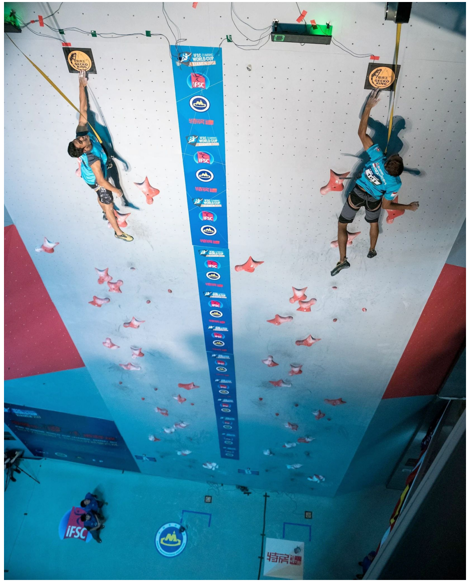
2019 - World Cup series used as Olympic Qualifying events

2020 - Olympics Pushed to 2021

Climbing Events

Speed

Tall, Fast, Known



Bouldering

Short, Difficult, Unknown

Leading

Tall, Difficult, Unknown

Climbing Events

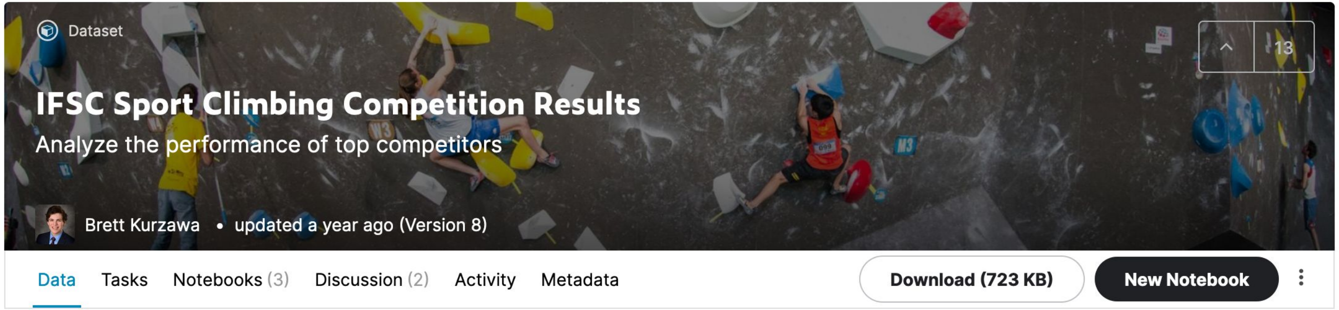
Speed		Bouldering		Leading
Ranking		Ranking		Ranking
5	×	2	×	1

=

Final Score: 10

Rank by final score

Data Used



https://www.kaggle.com/brkurzawa/ifsc-sport-climbing-competition-results

- Dataset scraped from IFSC website
- Results from all competitor from the 2018 and 2019 IFSC World Cup season
- Combined Competition plus Individual Competition (Speed, Bouldering, Lead)

Show Data

In [3]:
```python
cr_raw.tail()
```

Out[3]:

	Competition Title	Competition Date	FIRST	LAST	Nation	Rank	Qualification lead	Qualification speed	Qualification boulder	Final lead	Final speed	Final boulder	Points	Final Points	Category
874	IFSC Climbing World Championships Combined - I...	16 September 2018	Adam	ONDRA	CZE	2	48+2.	54.	12.	42+1.	5.	4T4z1162.	1296.0	10.0	combined
875	IFSC Climbing World Championships Combined - I...	16 September 2018	Jan	HOJER	GER	3	21 20.	13.	6.	26+6.	1.	3T4z864.	1560.0	24.0	combined
876	IFSC Climbing World Championships Combined - I...	16 September 2018	Kai	HARADA	JPN	4	40 9.	23.	1.	34+3.	4.	3T3z745.	207.0	60.0	combined
877	IFSC Climbing World Championships Combined - I...	16 September 2018	Tomoa	NARASAKI	JPN	5	37 11.	6.	4.	34 4.	6.	3T4z653.	264.0	72.0	combined
878	IFSC Climbing World Championships Combined - I...	16 September 2018	Kokoro	FUJII	JPN	6	23 19.	16.	3.	30+5.	3.	2T4z256.	912.0	90.0	combined

In [7]:
```python
# Here we have the top 5 women based on the average ranking in all disciplines multiplied together.
# We do okay here predicting the top two women climbers.
pred_aggs[pred_aggs.gender == 'F'][['first', 'last', 'rank', 'gender',
                                    'lead_avg_rank', 'boulder_avg_rank',
                                    'speed_avg_rank', 'avg_rank_multi'
                                   ]].sort_values('avg_rank_multi').head()
```

Out[7]:

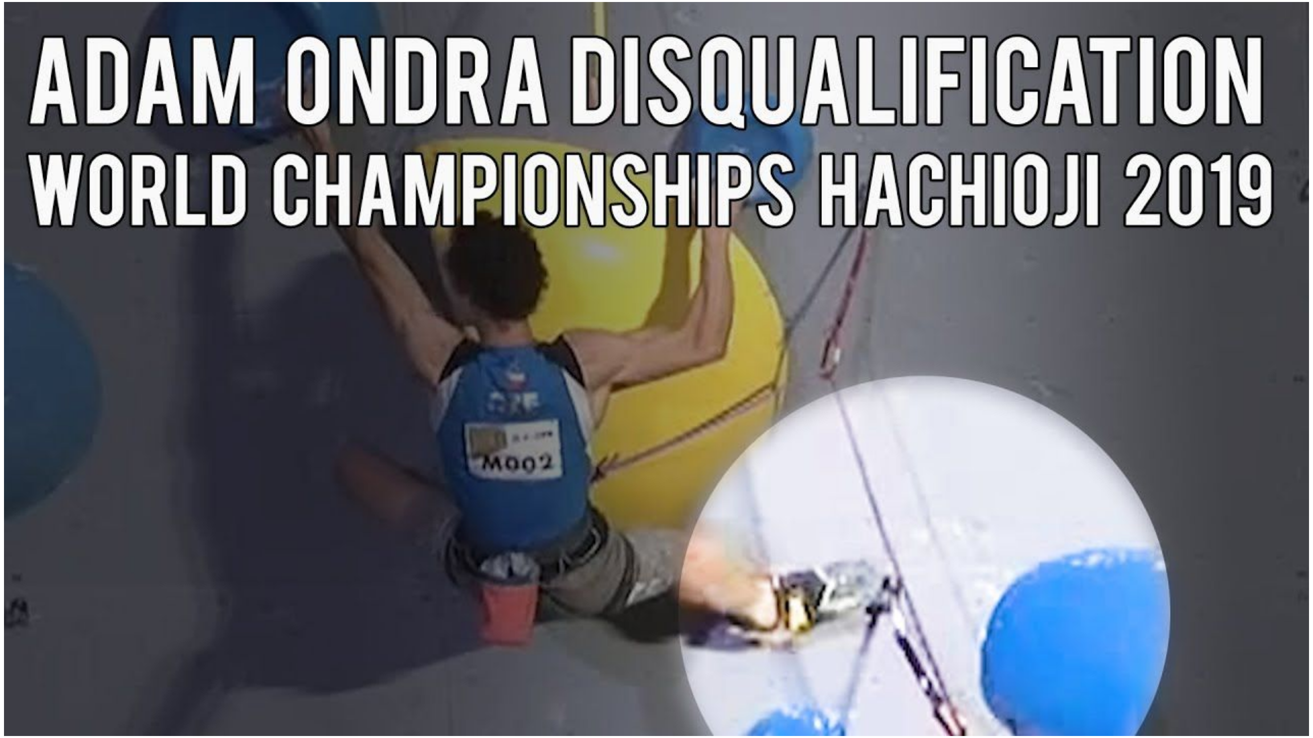
	first	last	rank	gender	lead_avg_rank	boulder_avg_rank	speed_avg_rank	avg_rank_multi
0	Janja	GARNBRET	1	F	2.181818	1.100000	34.888889	83.733333
1	Akiyo	NOGUCHI	2	F	5.375000	2.307692	35.461538	439.859467
4	Miho	NONAKA	5	F	14.142857	2.727273	24.111111	930.000000
9	Jessica	PILZ	10	F	3.000000	10.000000	41.777778	1253.333333
5	Ai	MORI	6	F	3.500000	17.000000	63.666667	3788.166667

In [9]:
```python
# Here we do the same for the men but our top prediction Adam Ondra got 18th.
# Slides about why this is and intro ranking metrics.
pred_aggs[pred_aggs.gender == 'M'][['first', 'last', 'rank', 'gender',
                                    'lead_avg_rank', 'boulder_avg_rank',
                                    'speed_avg_rank', 'avg_rank_multi'
                                   ]].sort_values('avg_rank_multi').head()
```

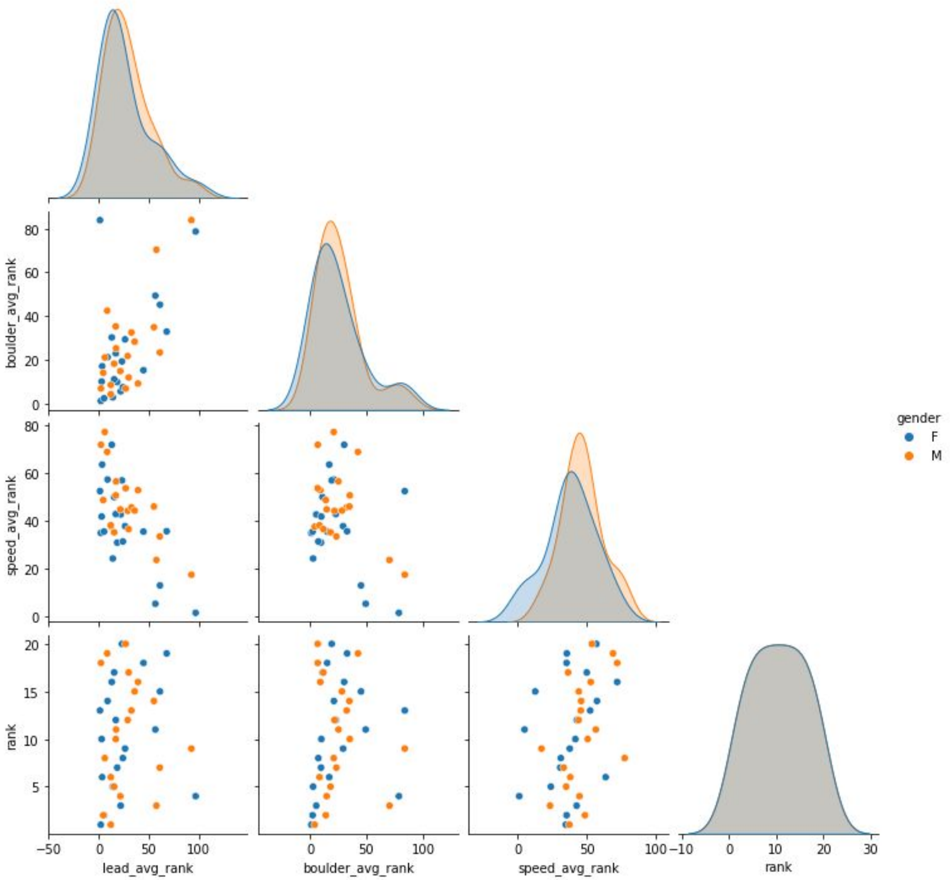
Out[9]:

	first	last	rank	gender	lead_avg_rank	boulder_avg_rank	speed_avg_rank	avg_rank_multi
37	Adam	ONDRA	18	M	2.333333	6.833333	72.000000	1148.000000
20	Tomoa	NARASAKI	1	M	12.166667	4.166667	37.500000	1901.041667
21	Jakob	SCHUBERT	2	M	4.600000	14.000000	48.727273	3138.036364
25	Kokoro	FUJII	6	M	12.142857	8.428571	38.083333	3897.712585
39	Jongwon	CHON	20	M	27.000000	6.769231	53.714286	9817.318681

Anomalous Events Happen



```
In [6]: # Here we can see each event average ranking relationship
sns.pairplot(pred_aggs[['lead_avg_rank', 'boulder_avg_rank', 'speed_avg_rank', 'rank', 'gender']]
             ,hue='gender'
             ,corner=True);
```



Ranking Metrics

Ranking Metrics

- Goal: Correctly pick the Olympic Medal winners
  - Want better climbers ranked higher
  - Care about ordering
- Normalized Discounted Cumulative Gain (nDCG):
  - Higher when relevant (rel_i) content is higher
  - Normalized by ideal ranking relevance score (IDCG)
  - Between 0 and 1
  - Can pick p to determine how far down the rankings you score
- Kendall Rank Coefficient:
  - If pairs are properly ordered (concordant) you get a higher score
  - Between -1 and 1

$$DCG_p = \sum_{i=1}^{p} \frac{rel_i}{\log_2(i+1)}$$

$$nDCG_p = \frac{DCG_p}{IDCG_p}$$

$$\tau = 2\,\frac{(\text{number concordant pairs}) - (\text{number of discordant pairs})}{n(n-1)}$$

Models + Metrics

In [28]:
```
final_results[final_results['gender'] == 'A']
```
Out[28]:

	pred_method	gender	ndcg_all	ndcg_3	kendall
0	avg_rank_multi_relev	A	0.937127	0.946710	0.306519
3	pred_LR	A	0.886157	0.665077	0.363667
6	pred_nn	A	0.899401	0.949985	0.093514
9	pred_t	A	0.916973	0.779545	0.266985

In [27]:
```
final_results[final_results['gender'] == 'F']
```
Out[27]:

	pred_method	gender	ndcg_all	ndcg_3	kendall
2	avg_rank_multi_relev	F	0.977016	0.974264	0.610526
5	pred_LR	F	0.869137	0.684434	0.326316
8	pred_nn	F	0.936308	0.819851	0.305263
11	pred_t	F	0.948581	0.982383	0.253444

In [26]:
```
final_results[final_results['gender'] == 'M']
```
Out[26]:

	pred_method	gender	ndcg_all	ndcg_3	kendall
1	avg_rank_multi_relev	M	0.807707	0.591601	0.052632
4	pred_LR	M	0.957975	0.912685	0.463158
7	pred_nn	M	0.826575	0.611525	-0.105263
10	pred_t	M	0.781653	0.432898	0.202911

Conclusions

- Simple can be good
- Using domain knowledge in data processing is key
  - Feature built off combine scoring method (multiplied event rankings)
  - Structuring test/train similar to model usage
  - Anomalies happen and will affect results
- Metrics should align to goals
- I’m gonna win all the Climbing Olympic betting pools

Thank You & Questions?