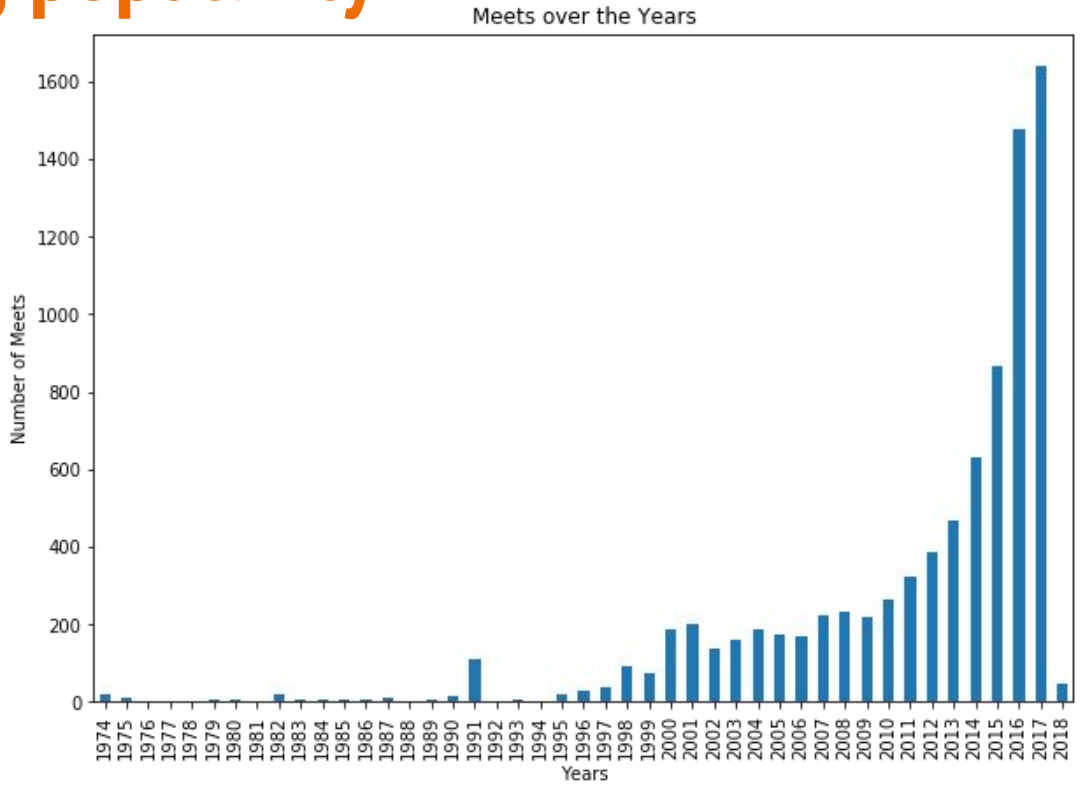
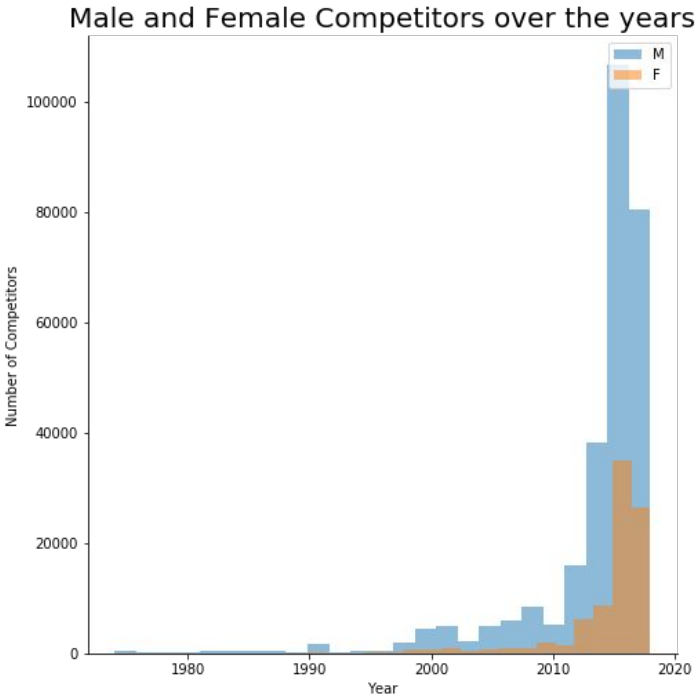
Strength Comes in all Sizes

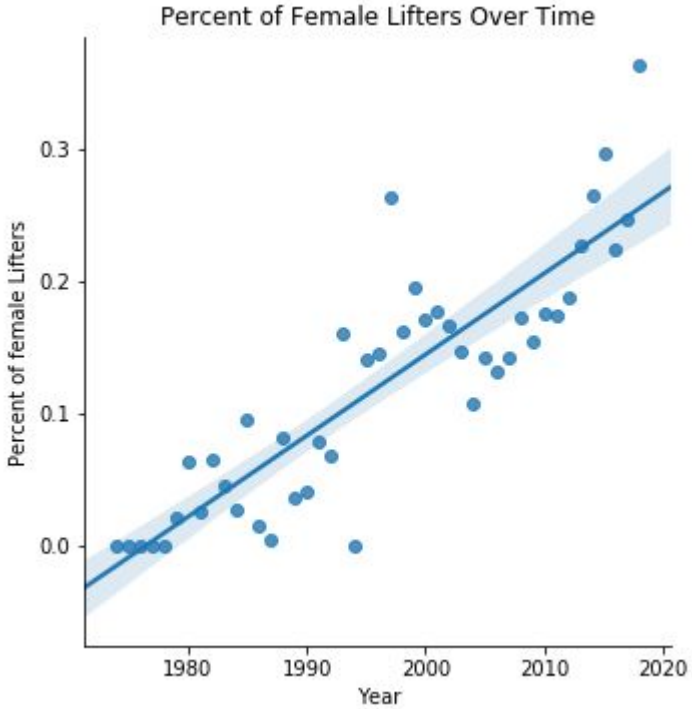
Or does it?

The rising popularity



Growth of the sport





What is Wilks score?

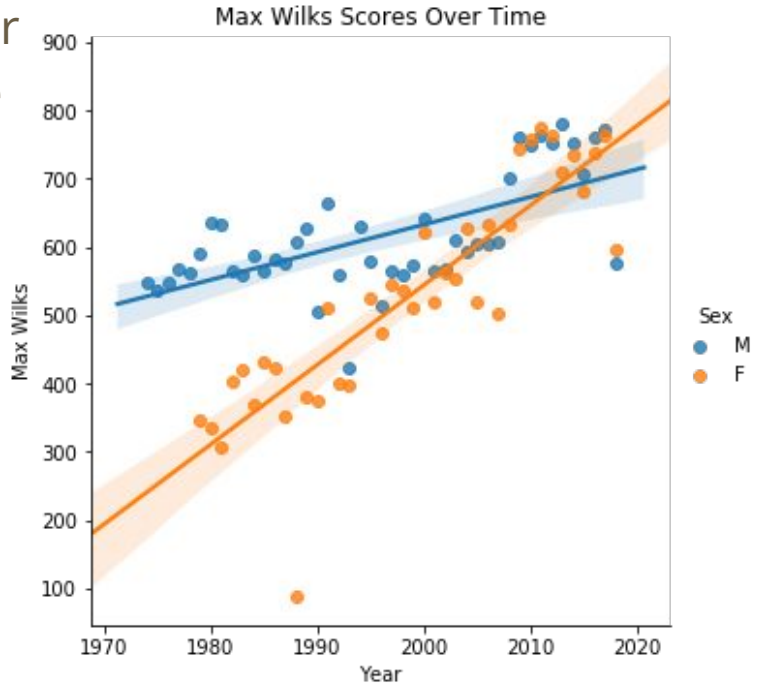
- Wilks coefficient multiplied against total lifted

$$Coeff = \frac{500}{a + bx + cx^2 + dx^3 + ex^4 + fx^5}$$

- Used to help normalize lifters scores, regardless of sex or weight
- Has known biases
  - Favorable towards intermediate women weight classes on deadlift
  - Unfavorable towards heavy men and women on squat

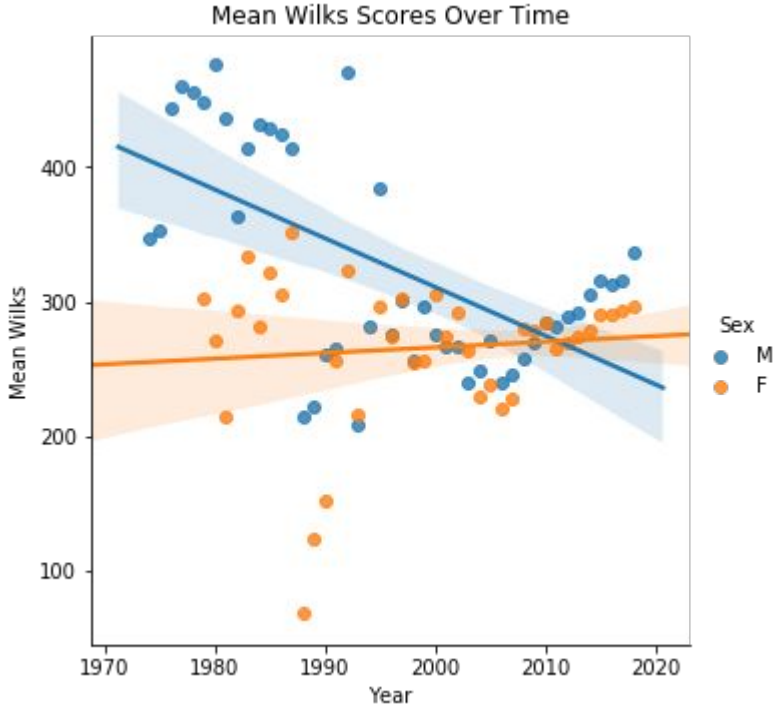
Max WIlk's Scores Over Time

- Men have always been considered stronger
- Current highest Wilks score held by female
- The gap in strength is closing



Average Wilks Scores over time

- Men’s scores are dropping
- Women’s scores increasing slightly



Predicting Performance

- Wilks most quantifiable data to be predicted
- With biases, still good indicator of performance
- Can the wilks score be predicted?

Biggest Predictors

- Competing in multiple lifts most indicative
- Following close behind, competing in just bench
- M(male) 8th out of 11

	Feature	Importance
9	Multiple	0.469072
7	Bench	0.317297
2	Num_meets	0.048531
8	Deadlift	0.042426
0	Year	0.029027
5	Wraps	0.020695
1	BodyweightKg	0.019135
3	Raw	0.018376
6	M	0.018351
4	Single-ply	0.015491
10	Squat	0.001600

The models

- Three linear models performed almost identically
- Random Forest predicts on unseen data almost as well as training data
- K nearest neighbors performed the best overall
    - Lifters perform closest to similar lifters

	r2_train	r2_test	mse_train	mse_test
LinearRegression	0.615987	0.615679	4948.4	4958.63
Lasso	0.615987	0.615682	4948.39	4958.59
Ridge	0.615987	0.615682	4948.39	4958.59
RandomForestRegressor	0.685127	0.681784	4057.46	4105.73
KNeighborsRegressor	0.701918	0.686281	3841.09	4047.7

Conclusions

- The gap between in strength between the sexes is closing
- Lifts competed in have biggest influence on score
- People perform like others
- Even with records being broken regularly,