

A decorative border in a light gold color frames the entire image. It consists of a stepped rectangular line with 'X' marks at the top and bottom center. On the left and right sides, there are stylized, overlapping geometric shapes resembling fan-like structures or abstract flowers.

OSCAR SO PREDICTABLE

Marta Lew
February 23, 2021

A decorative geometric ornament at the top center of the slide, featuring a central diamond shape with radiating lines and fan-like patterns on either side.

PROBLEM STATEMENT

Examine trends in the entertainment industry and attempt to predict Academy Award favorites based on information such as: release date, plot summary, genre, production company, budget, opening WE results, and history of Oscars among a movie's cast & crew.

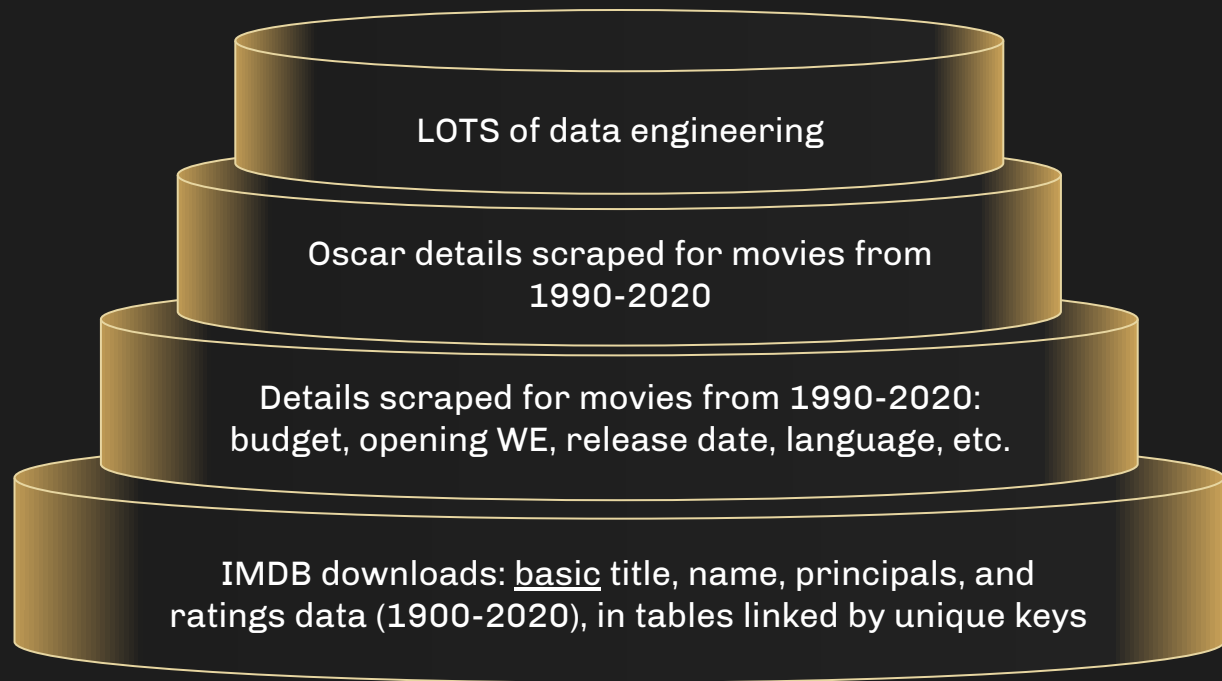
What's the big deal with the Oscars?

The Academy Awards mean much more than a celebration of talent:

- ◇ ~\$149 M in revenue for ABC - the exclusive home of the Oscars since 1997
- ◇ ~\$130 M in economic boost that LA sees from hosting the Oscars
- ◇ ~\$100 M+ is spent each year on awards season lobbying
- ◇ ~\$15-20 M average increase in box office sales for a Best Picture winner
- ◇ 20% is the boost in pay that a Best Actor/Actress can expect on their next film
- ◇ \$150-225k goes to every nominee in party favors
- ◇ ...

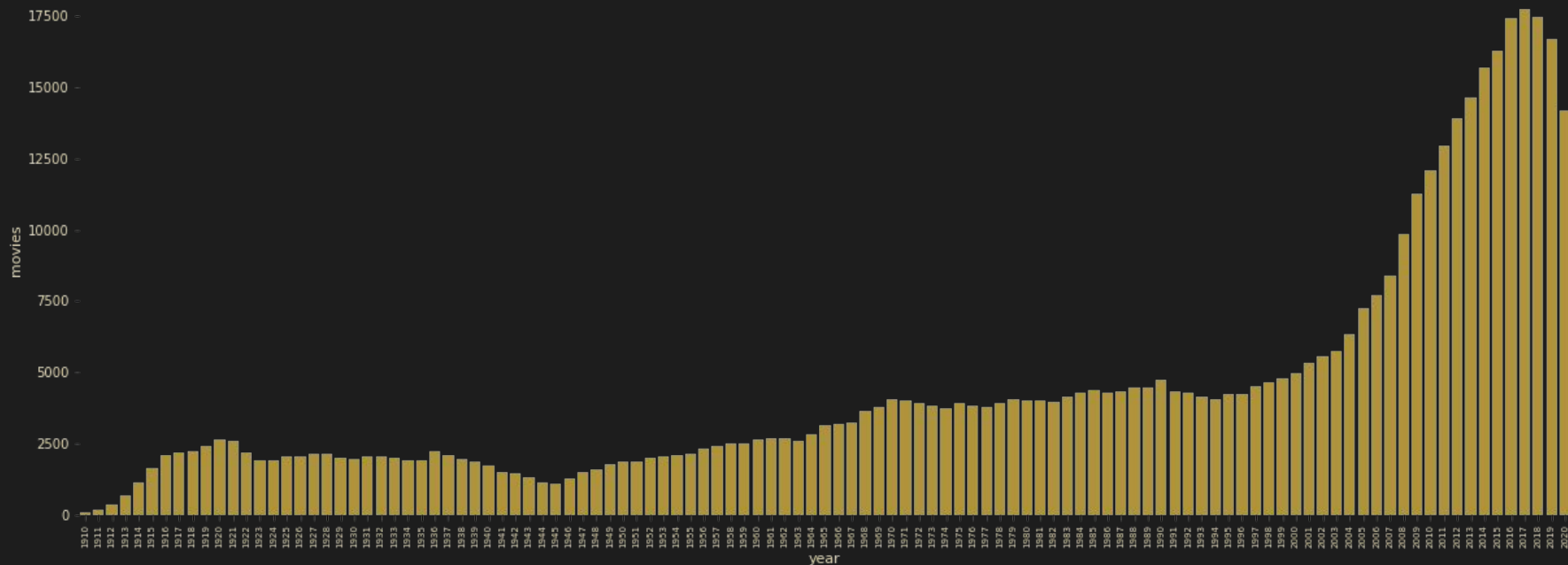


Big deal - big datasets



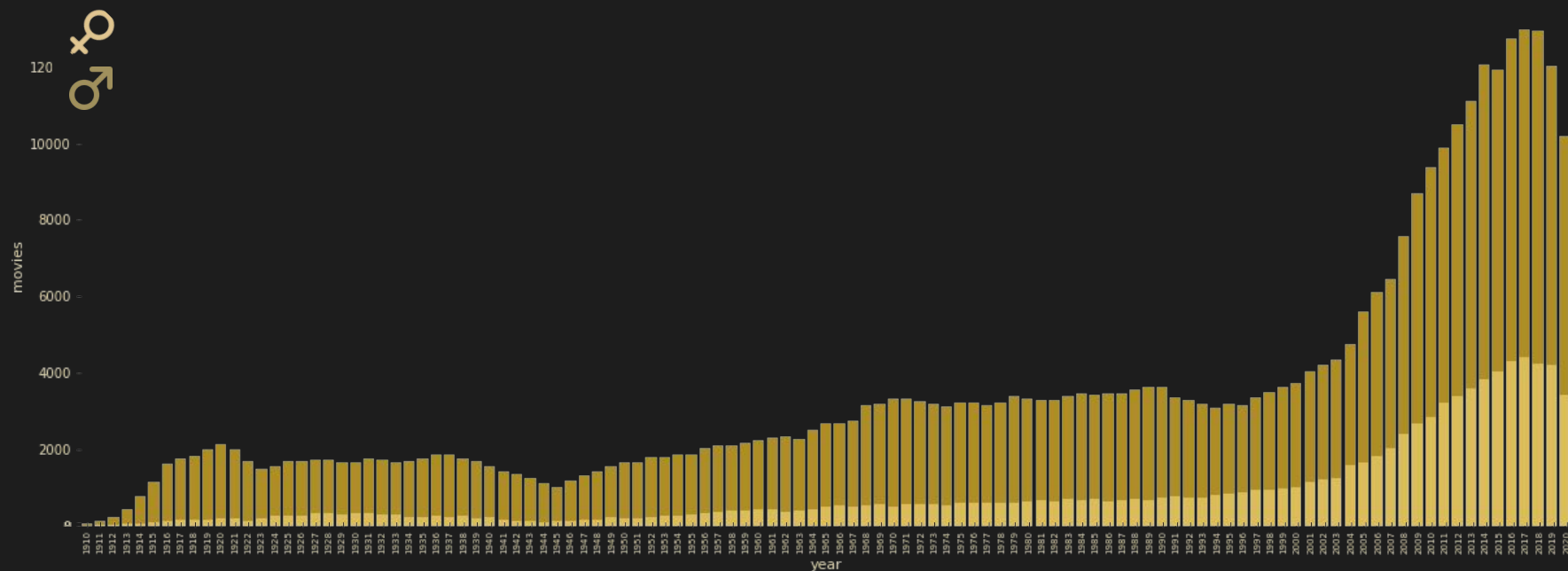
The movies, in numbers

Number of movies produced worldwide, by year



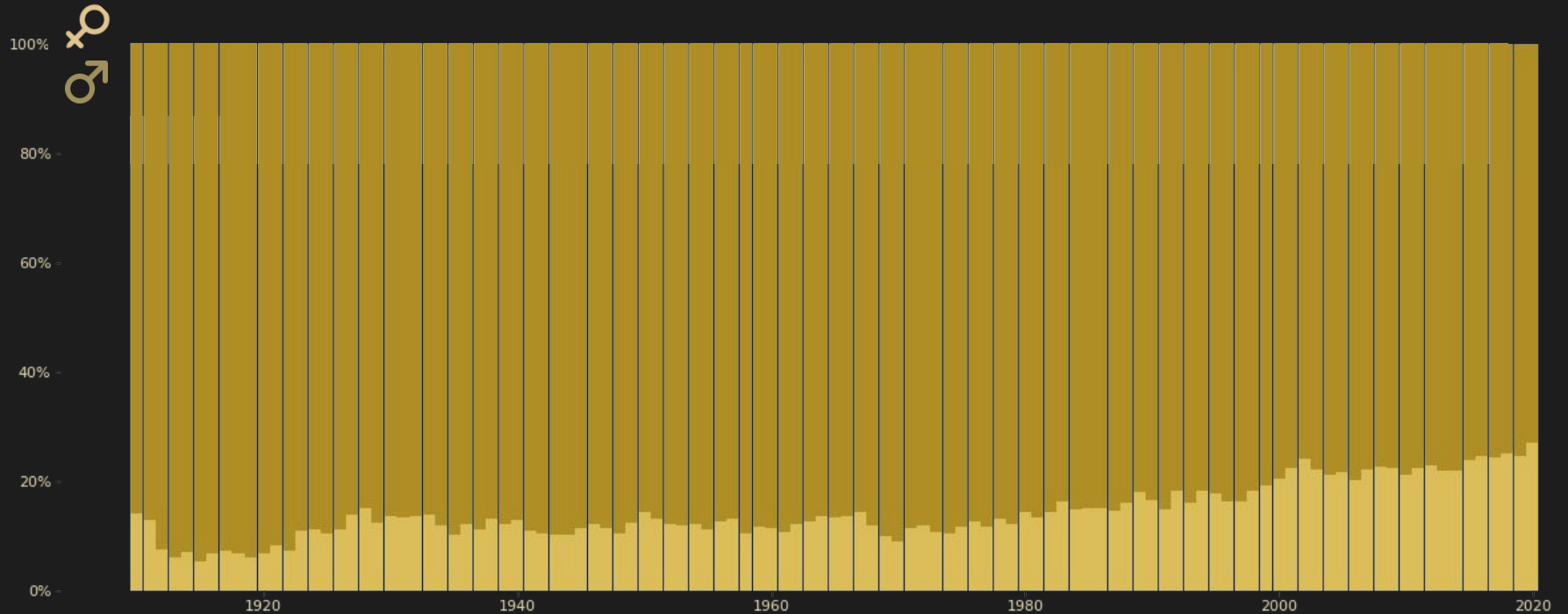
The movies, in numbers

Number of movies made each year, by gender of director



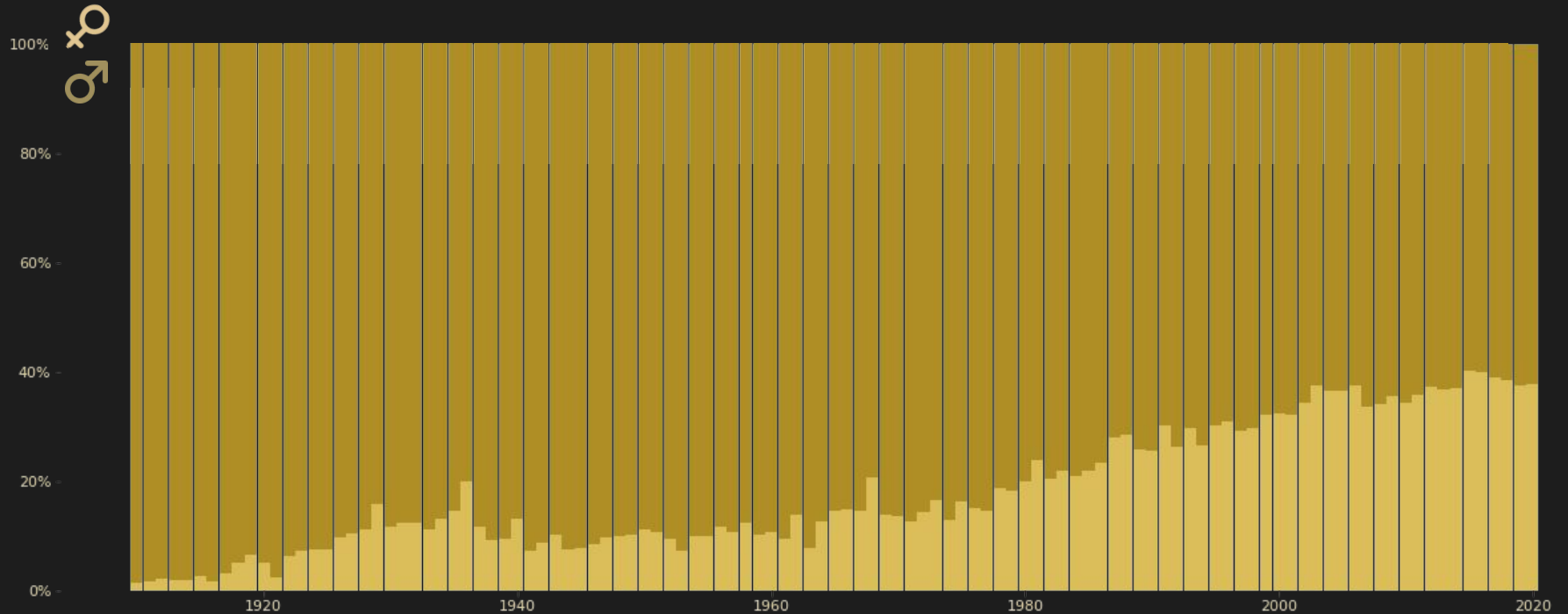
Industry diversity over time

Proportion of movies made each year by female and male directors



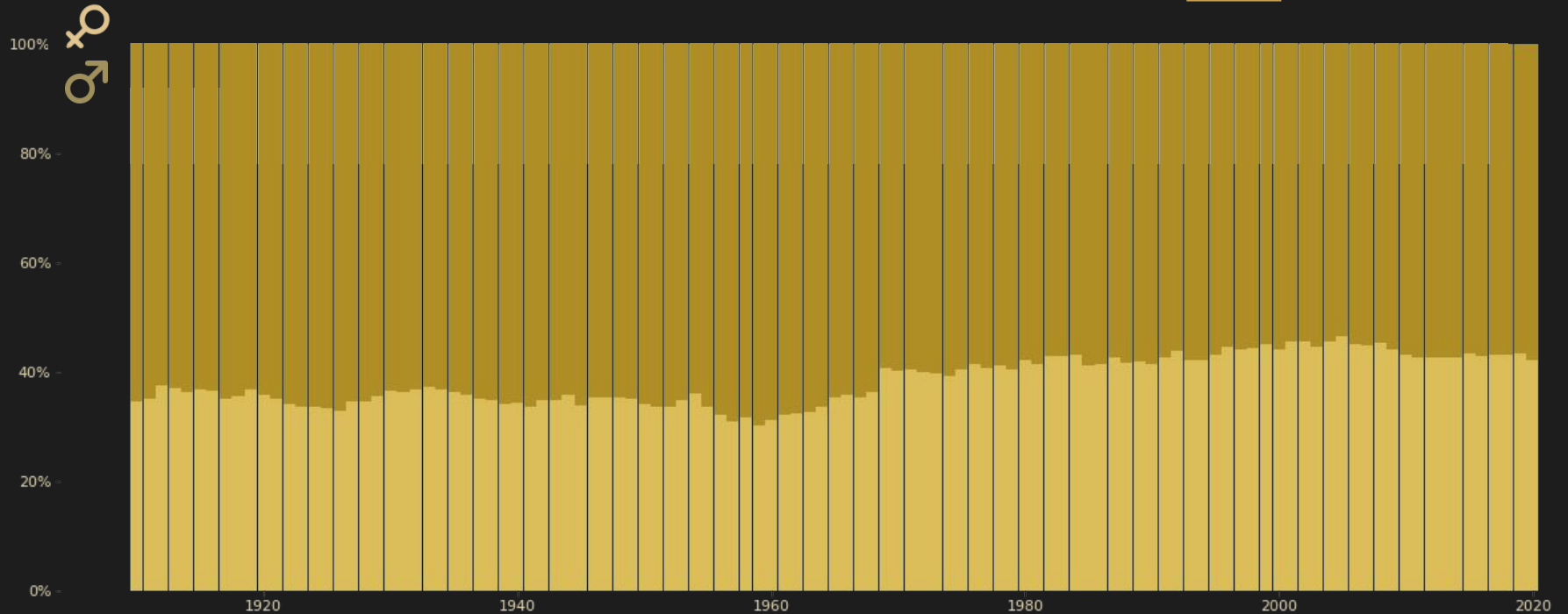
Industry diversity over time

Proportion of movies made each year by female and male producers



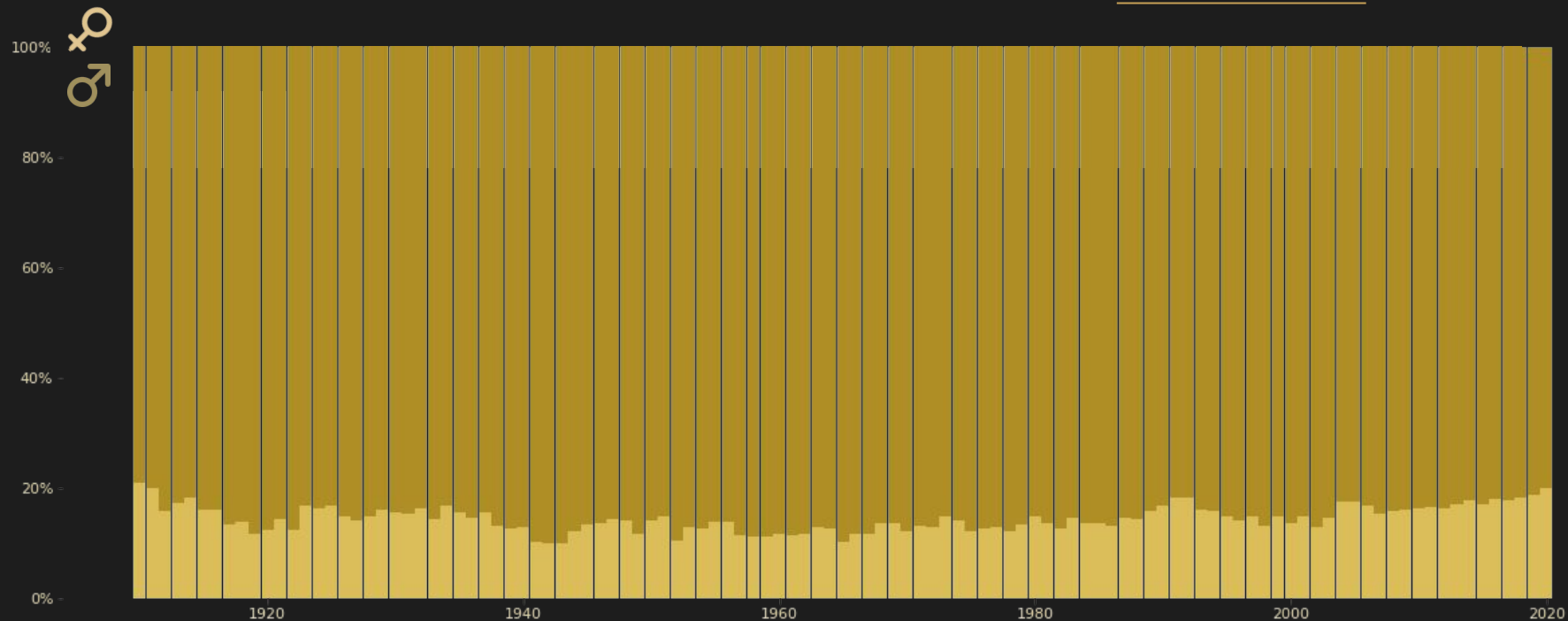
Industry diversity over time

Proportion of movies made each year by female and male actors



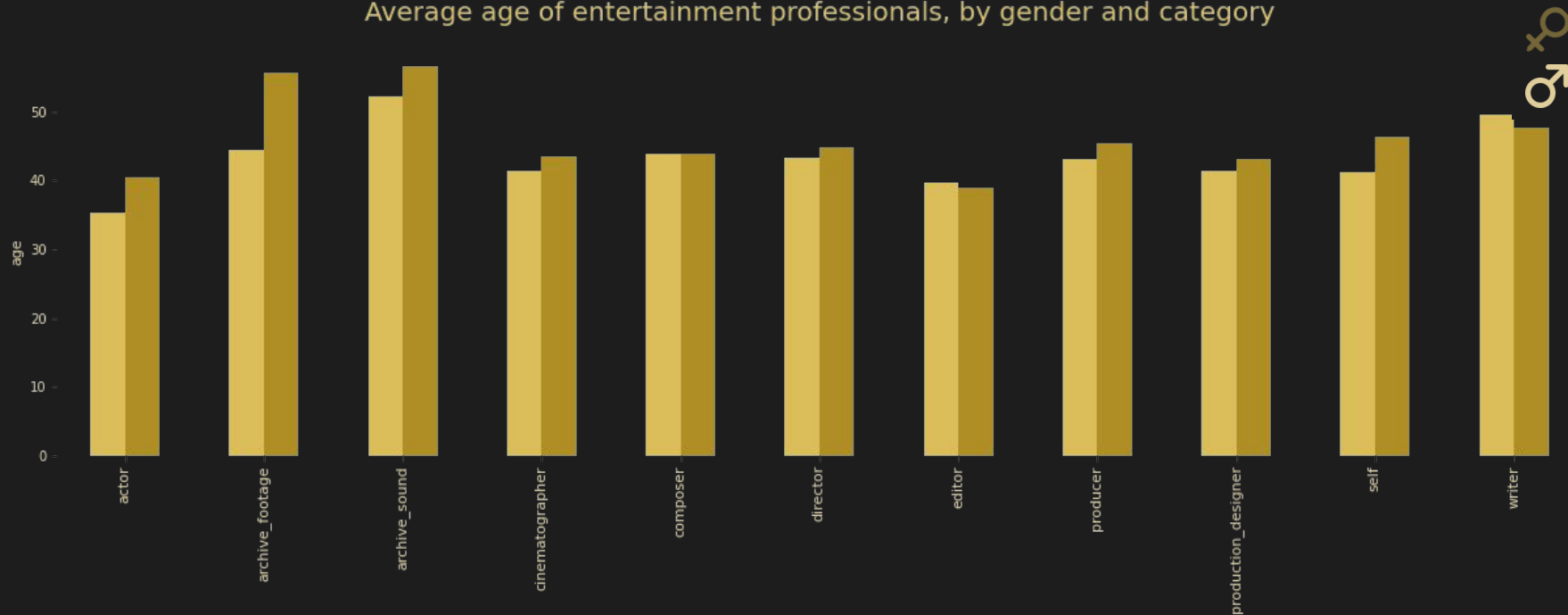
Industry diversity over time

Proportion of movies made each year by female and male cinematographers



Age differences

Average age of entertainment professionals, by gender and category



Back to the Oscars



These aren't actually Oscar statuettes but close enough

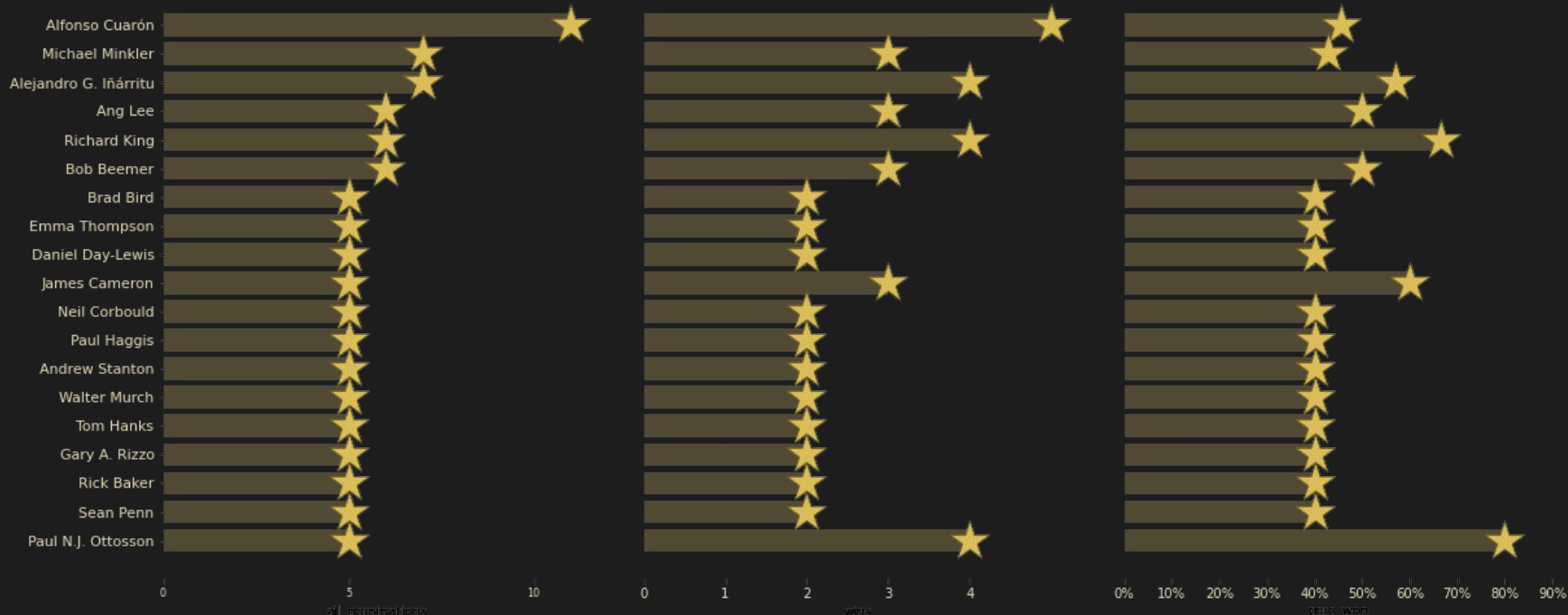
The Academy is very generous to some...

Academy Sweethearts

Nominations received from 1990 to 2020

Awards received from 1990 to 2020

Awards received from 1990 to 2020
as % of received nominations



★ Win/-s

ACADEMY LOVE

Academy loving its sweethearts (1910-2020)

The Academy Loving Its Sweethearts (1990-2020)

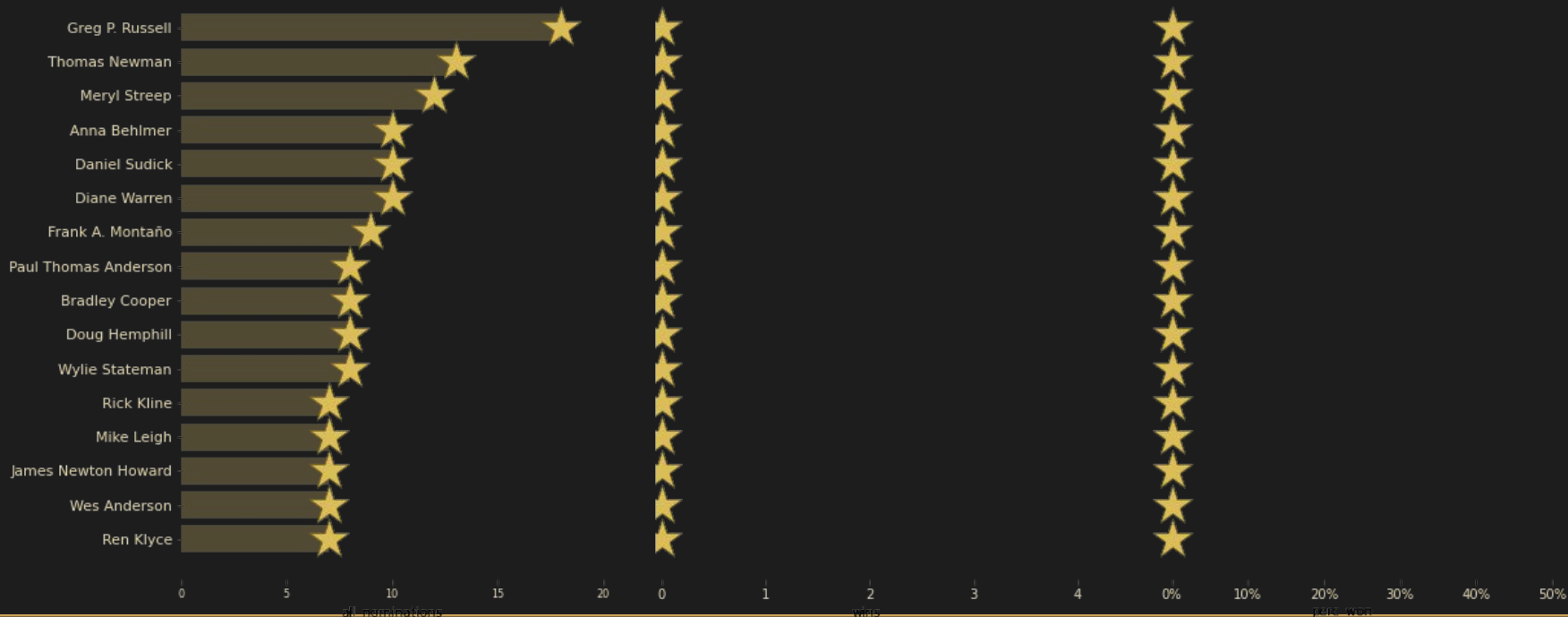
The Academy is not generous to all

Academy Sweetsnubs

Nominations received from 1990 to 2020

Awards received from 1990 to 2020

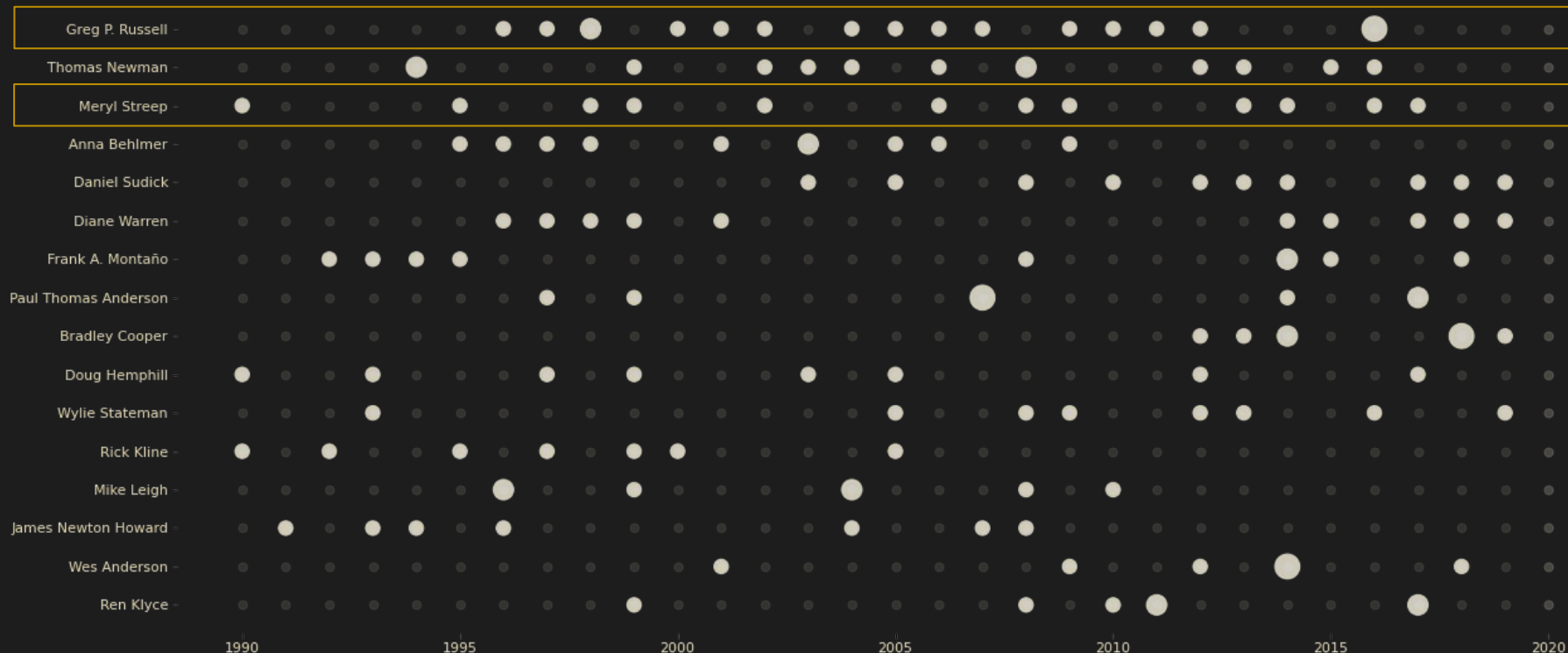
Awards received from 1990 to 2020
as % of received nominations



- Nomination/-s
- ★ Win/-s

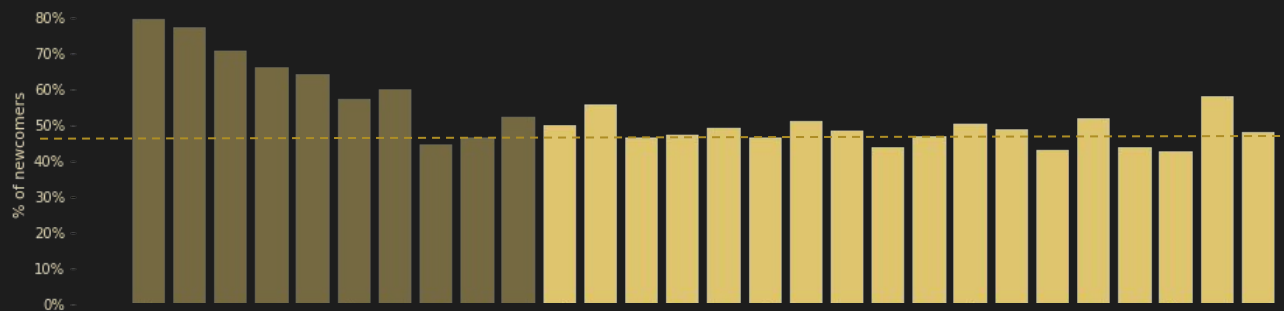
ACADEMY LOVE... of snubbing

Academy snubbing its sweetsnubs (1910-2020)

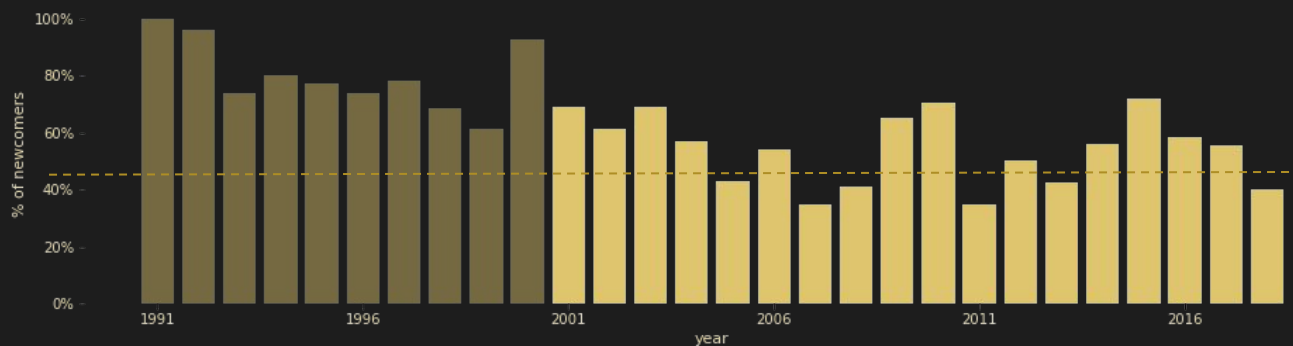


Can newcomers break in?

Number # of Oscar newcomers nominated by the Academy as % of all nominees, by year (1990-2020)



Number # of Oscar newcomers awarded by the Academy as % of all winners, by year (1990-2020)

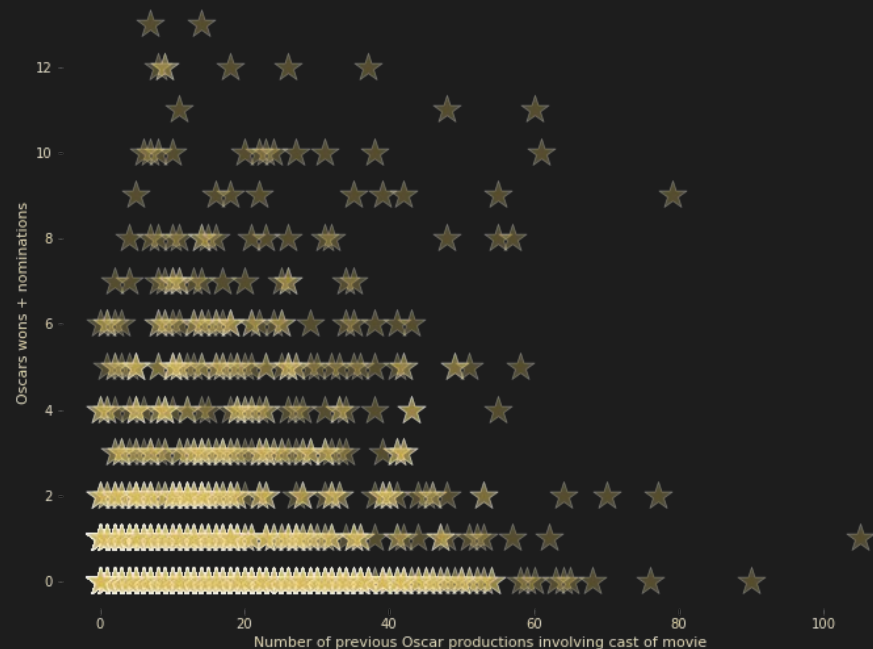


Once a winner, always a winner?

Oscar wins & nominations vs Oscar-density of cast and crew,
by movie (1995-2020)

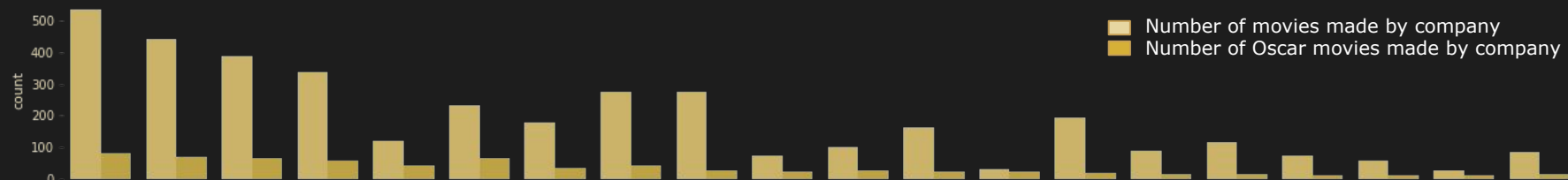


Oscar wins & nominations vs Oscar-density of cast and crew,
by movie (1995-2020)

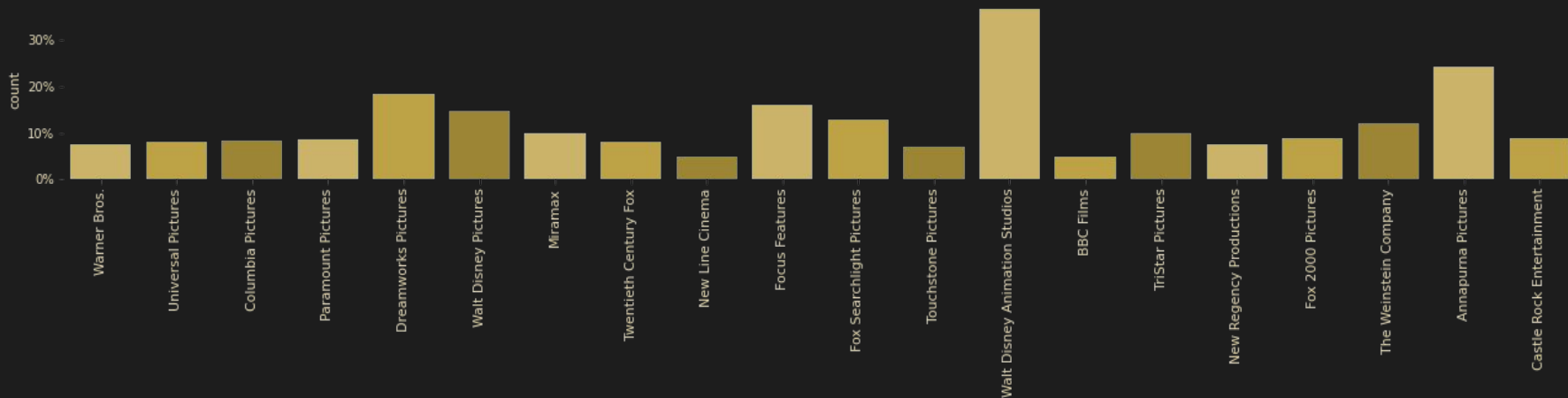


Oscar luck by studio

Number of movies / Oscar movies made by key production companies

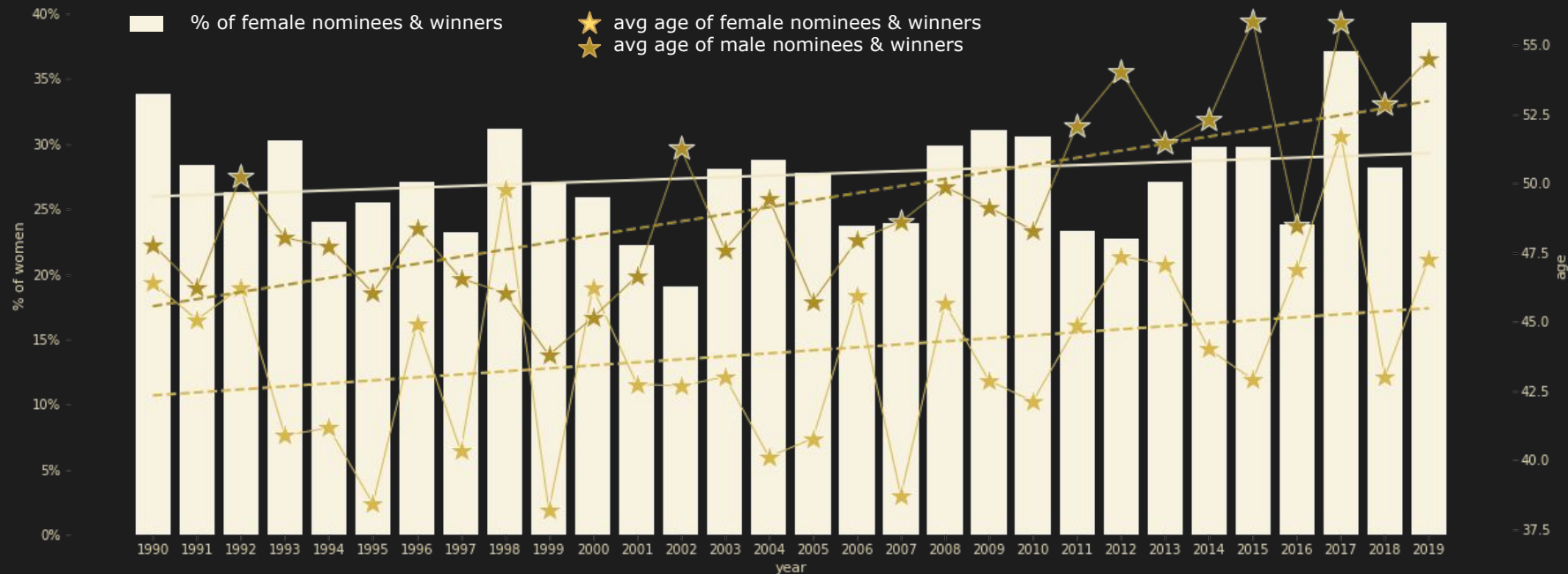


Oscar movies made by key production companies as % of all movies made by each company



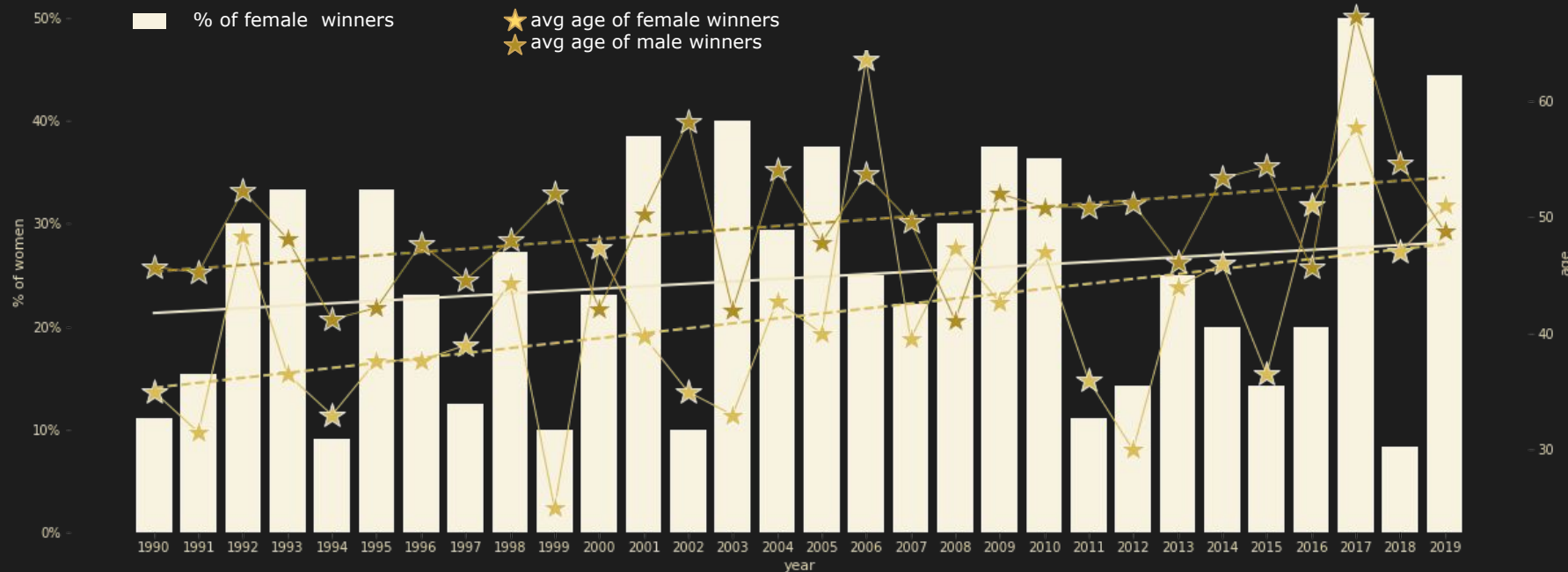
Academy's support of diversity

Average age of Academy Award nominees and recipients



Academy's support of diversity

Average age of Academy Award recipients





THE MAIN — ACT: — PREDICTIONS

The magnitude and complexity of the task

175,693
movies made
between 1990
and 2019



1,197
movies recognized by
the Academy during
this period

0.7%



237
movies which
received Academy
Awards

0.1%

Four questions - four predictors

Which movies are most
likely to receive a
statuette?

AWARD

Which movies have the
best chances of winning
THE award?

BEST PIC

O1

O2

O3

O4

NOMINATION

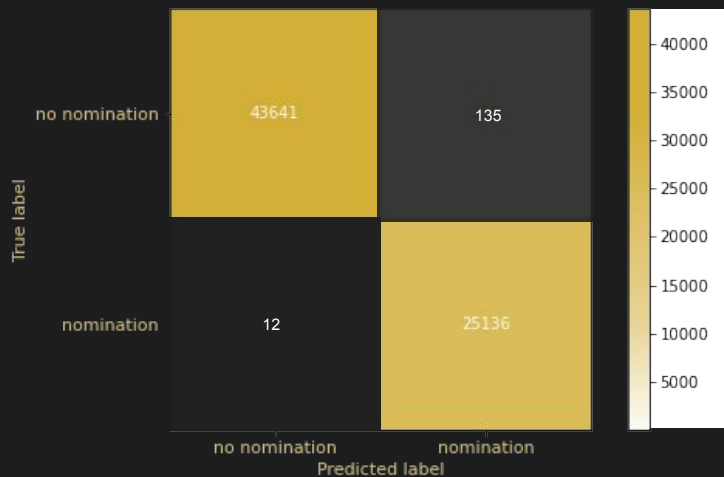
Which movies will receive
an Oscar nomination, in
any category?

BEST PIC NOMINATION

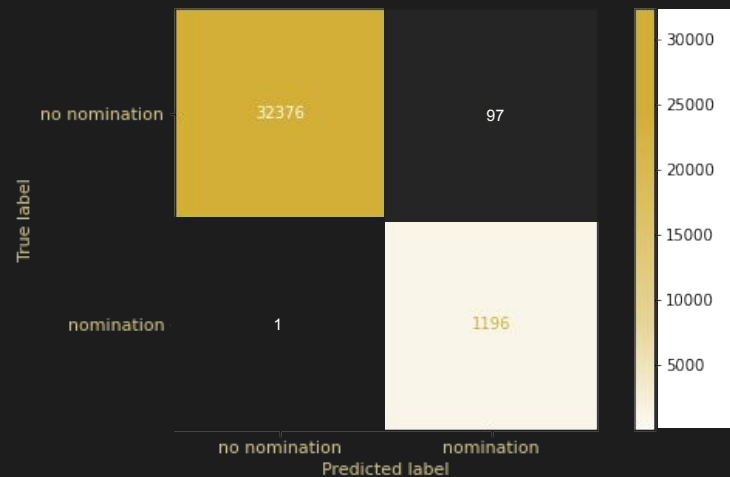
Which movies will the
Academy nominate for
Best Motion Picture?

Predicting movies to be nominated for an Oscar

Confusion matrix with duplicates from bootstrapping



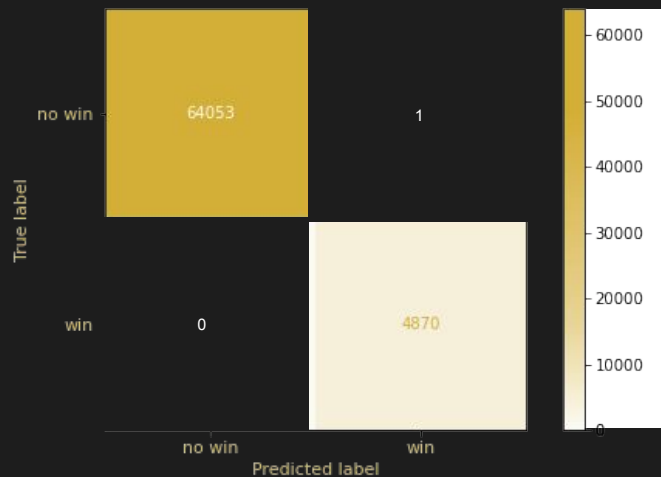
Confusion matrix without duplicates from bootstrapping



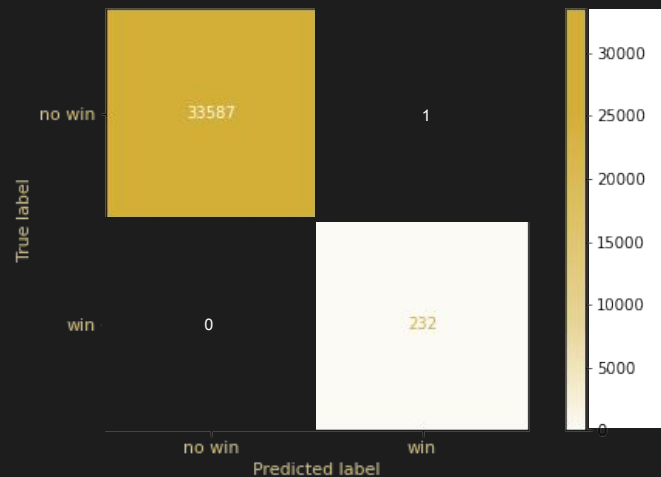
- ★ ACCURACY: 0.9986 on train data, 0.9974 on test data. 0.9921 in 5-fold cross-validation
- ★ RECALL: 0.9992 on train data, 0.9995 on test data. 0.9992 in 5-fold cross-validation

Predicting movies to receive an Oscar

Confusion matrix with duplicates from bootstrapping



Confusion matrix without duplicates from bootstrapping



- ★ ACCURACY: 1.0 on train data, 1.0 on test data. 1.0 in 5-fold cross-validation
- ★ RECALL: 1.0 on train data 1.0 on test data. 1.0 in 5-fold cross-validation

Predicting movies to be nominated for Best Picture

Confusion matrix with duplicates from bootstrapping



Confusion matrix without duplicates from bootstrapping



- ★ ACCURACY: 1.0 on train data, 0.9974 on test data. 0.9999 in 5-fold cross-validation
- ★ RECALL: 1.0 on train data, 0.9995 on test data. 1.0 in 5-fold cross-validation

Predicting movies to win Best Picture

Confusion matrix with duplicates from bootstrapping



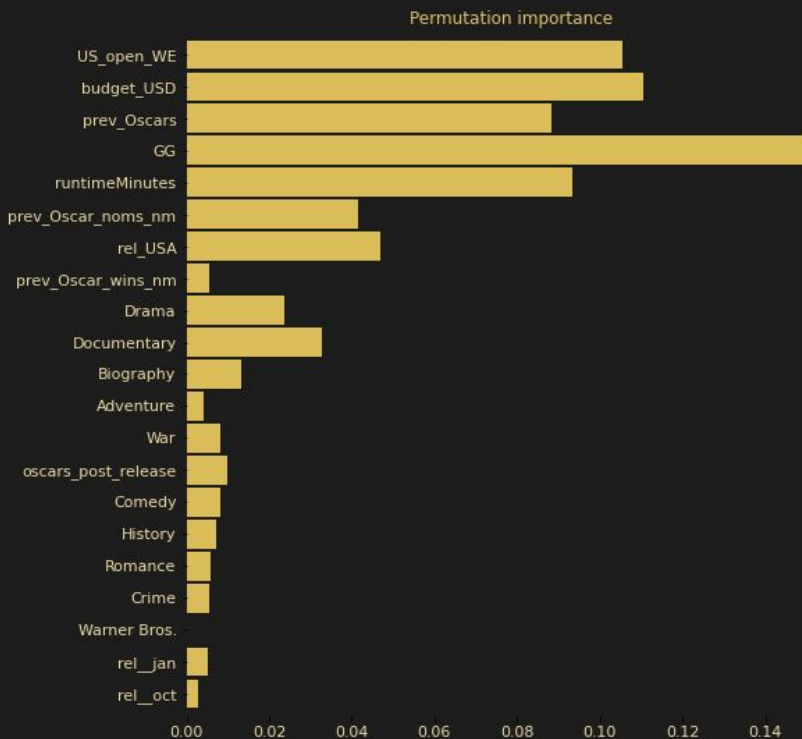
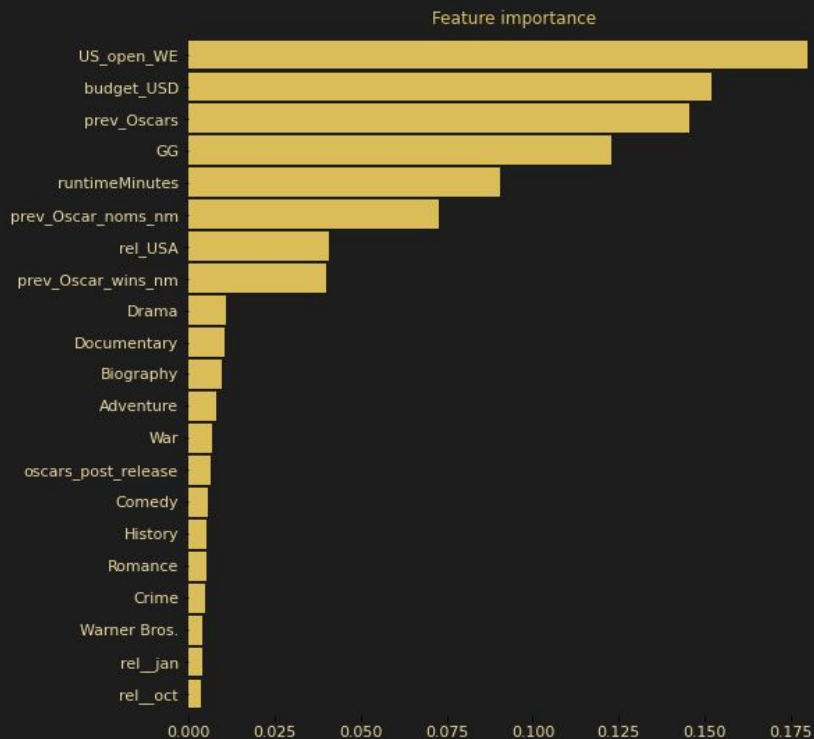
Confusion matrix without duplicates from bootstrapping



- ★ ACCURACY: 1.0 on train data, 0.9974 on test data. 1.0 in 5-fold cross-validation
- ★ RECALL: 1.0 on train data, 0.9995 on test data. 1.0 in 5-fold cross-validation

How does it know???

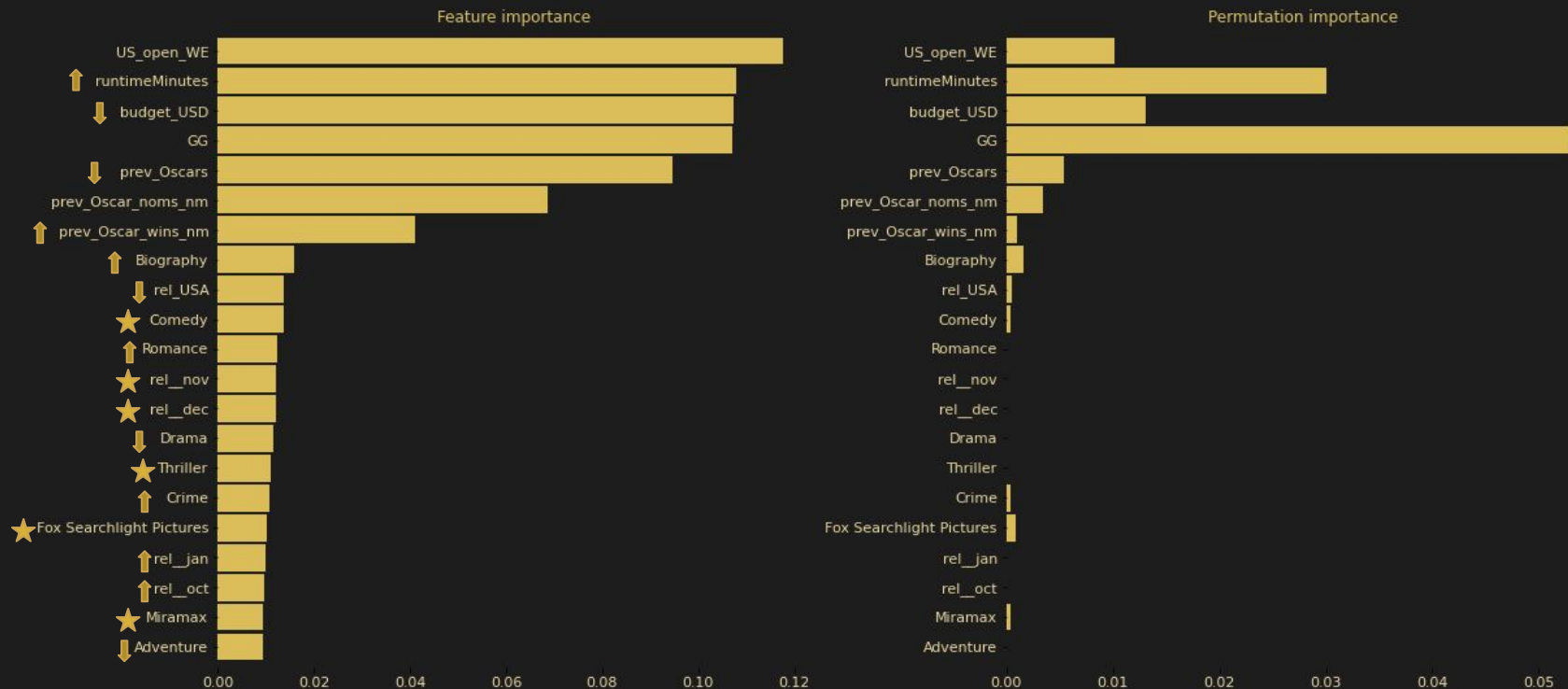
Feature importance in a Random Forest Classifier of Oscar contenders



How does it know???

Compared to all
contenders

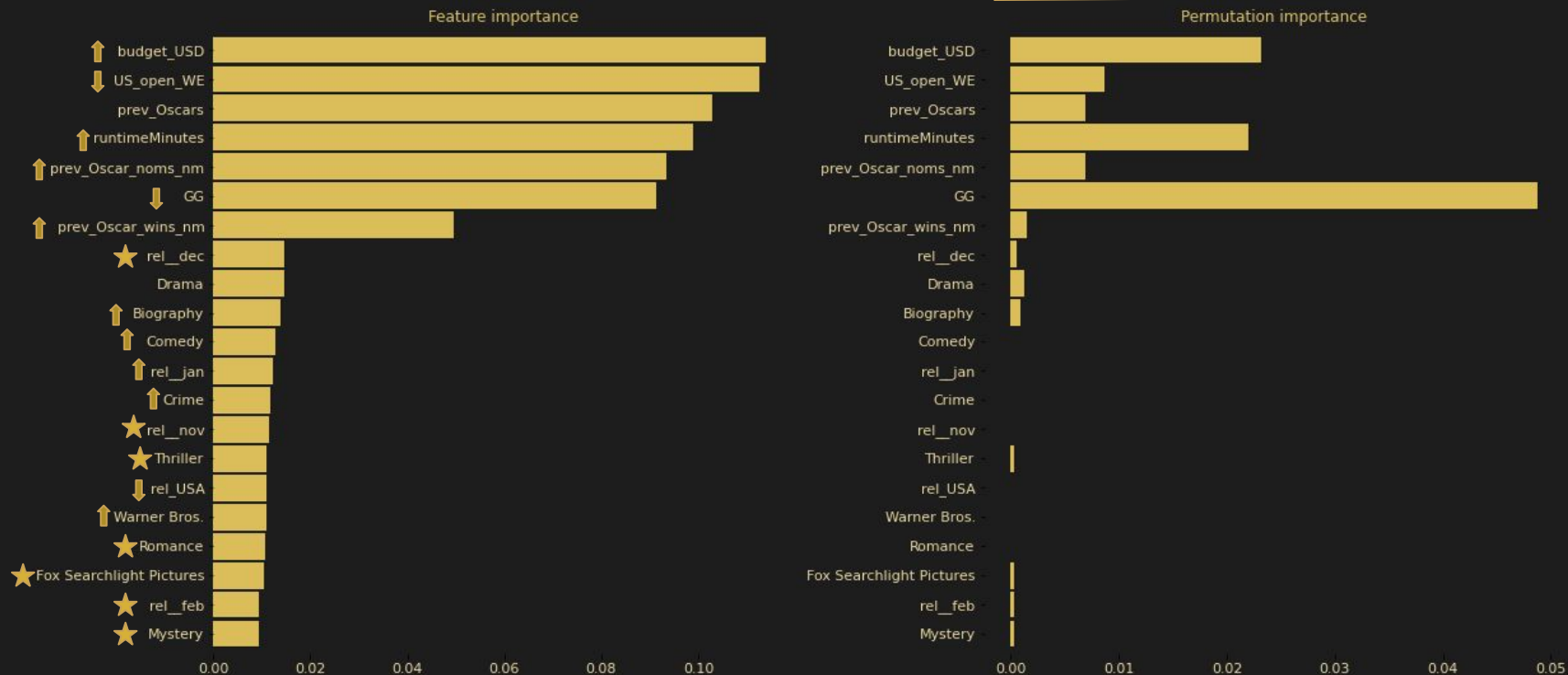
Feature importance in a Random Forest Classifier of Oscar winners



How does it know???

Compared to all
contenders

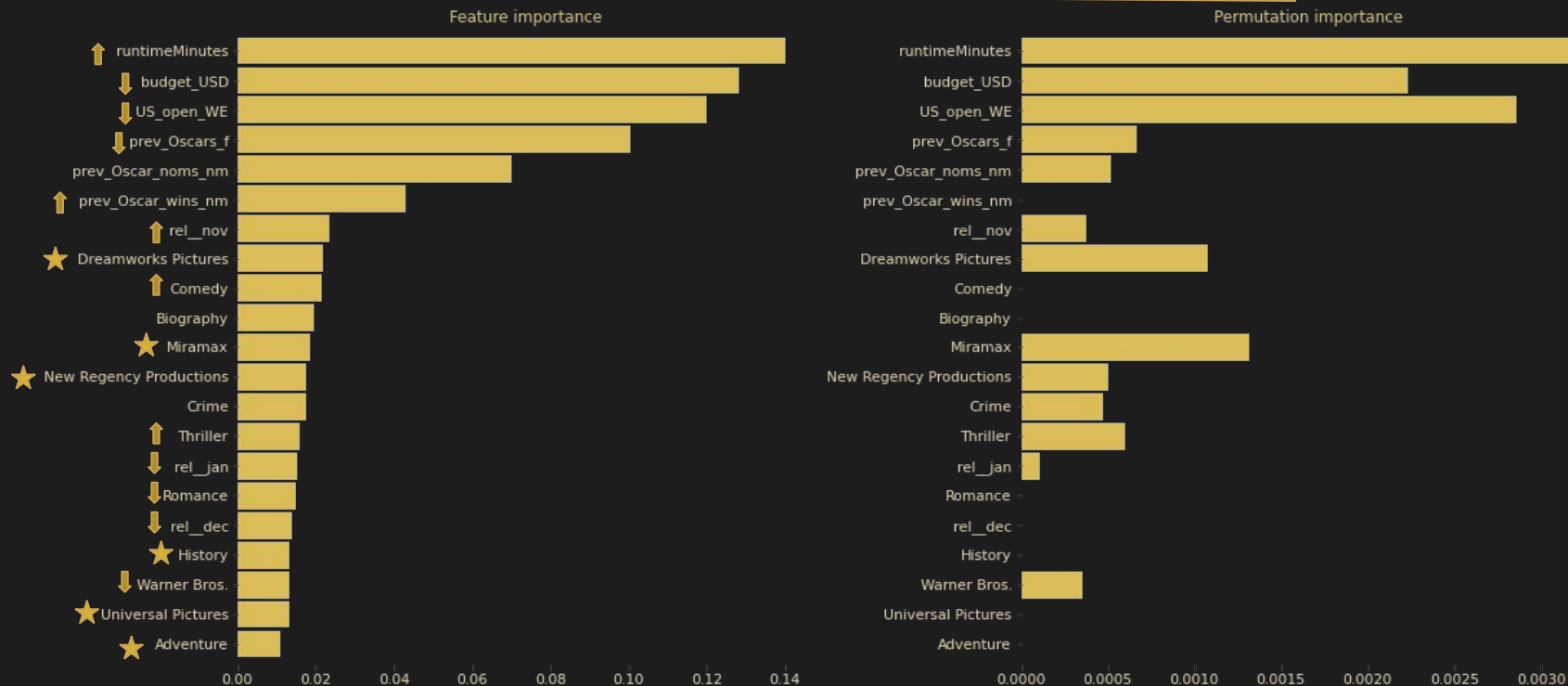
Feature importance in a Random Forest Classifier of Best Picture contenders



How does it know???

Compared to BP
contenders

Feature importance in a Random Forest Classifier of Best Picture winners



PREDICTIONS: 2021 Oscar Contenders

Title	Genres	Runtime	Produced by	Release date	Budget (USD)	US open WE	prev_Oscars
The Call of the Wild	Adventure,Drama,Family	100	20th Century Studios, 3 Arts Entertainment, Halon Entertainment	21 February 2020	\$135,000,000	\$24,791,624	26
Queen Elizabeth II: End of A Reign	Biography, Documentary	60	Not listed	14 July 2020	n/a	n/a	0
Ammonite	Biography, Drama, Romance	120	See-Saw Films, British Film Institute (BFI), BBC Films	13 November 2020	n/a	\$87,552	23
Onward	Adventure, Animation, Comedy	102	Walt Disney Pictures, Pixar Animation Studios	6 March 2020	n/a	\$39,119,861	16
The Eight Hundred	Action, Drama, History	149	Beijing Diqi Yinxiang Entertainment, Huayi Brothers Media	28 August 2020	\$80,000,000	\$118,161	2
Tesla	Biography, Drama	102	Millennium Media, Passage Pictures (II), Campbell Grobman Films	21 August 2020	\$5,000,000	\$42,165	7
Dolittle	Adventure, Comedy, Family	101	Universal Pictures, Dentsu, Perfect World Pictures	17 January 2020	\$175,000,000	\$21,844,045	31
The Croods: A New Age	Adventure, Animation, Comedy	95	DreamWorks Animation	25 November 2020	n/a	\$9,724,200	23
The Prom	Comedy, Drama, Musical	130	Dramatic Forces, Storykey, Netflix	11 December 2020	n/a	n/a	31
Hillbilly Elegy	Drama	116	Imagine Entertainment, Netflix	24 November 2020	n/a	n/a	41

PREDICTIONS: 2021 Oscar Winners

Title	Genres	Runtime	Produced by	Release date	Budget (USD)	US open WE	prev_Oscars
Wendy	Drama, Fantasy	111	Fox Searchlight Pictures, TSG Entertainment, Department of Motion Pictures	17 April 2020	\$6,000,000	\$26,214	4
Tenet	Action, Sci-Fi, Thriller	150	Warner Bros., Syncopy	3 September 2020	\$205,000,000	\$20,200,000	19
Mank	Biography, Comedy, Drama	131	Netflix	4 December 2020	\$10,000	n/a	34
The Invisible Man	Drama, Horror, Mystery	124	Universal Pictures, Blumhouse Productions, Goalpost Pictures	28 February 2020	\$7,000,000	\$28,205,665	6
Greyhound	Action, Drama, History	91	Sony Pictures Entertainment (SPE), Stage 6 Films, Bron Creative	10 July 2020	\$50,300,000	n/a	30
The Trial of the Chicago 7	Drama, History, Thriller	129	Dreamworks Pictures, Amblin Partners, Aperture Media Partners	16 October 2020	\$10,000	n/a	20
Ammonite	Biography, Drama, Romance	120	See-Saw Films, British Film Institute (BFI), BBC Films	13 November 2020	\$10,000	\$87,552	23
The Prom	Comedy, Drama, Musical	130	Dramatic Forces, Storykey, Netflix	11 December 2020	\$10,000		31
News of the World	Action, Adventure, Drama	118	Perfect World Pictures, Playtone, Pretty Pictures	25 December 2020	\$10,000	\$2,250,430	33
The Glorias	Biography, Drama, History	139	Artemis Rising Foundation, Page Fifty-Four Pictures, The Glorias	30 September 2020	\$10,000	n/a	27

PREDICTIONS: 2021 Best Picture Nominees

Title	Genres	Runtime	Produced by	Release date	Budget (USD)	US open WE	prev_Oscars
Mank	Biography, Comedy, Drama	131	Netflix	4 December 2020	\$10,000	n/a	34
Hillbilly Elegy	Drama	116	Imagine Entertainment, Netflix	24 November 2020	\$10,000	n/a	41
The Invisible Man	Drama, Horror, Mystery	124	Universal Pictures, Blumhouse Productions, Goalpost Pictures	28 February 2020	\$7,000,000	\$28,205,665	6
News of the World	Action, Adventure, Drama	118	Perfect World Pictures, Playtone, Pretty Pictures	25 December 2020	\$10,000	\$2,250,430	33
Wendy	Drama, Fantasy	111	Fox Searchlight Pictures, TSG Entertainment, Department of Motion Pictures	17 April 2020	\$6,000,000	\$26,214	4
The Trial of the Chicago 7	Drama, History, Thriller	129	Dreamworks Pictures, Amblin Partners, Aperture Media Partners	16 October 2020	\$10,000	n/a	20
The Prom	Comedy, Drama, Musical	130	Dramatic Forces, Storykey, Netflix	11 December 2020	\$10,000	n/a	31
Shikara	Drama, Romance	118	Vinod Chopra Productions	7 February 2020	\$3,836,000	n/a	4
Freaky	Comedy, Horror, Thriller	102	Blumhouse Productions, Divide/Conquer	13 November 2020	\$5,000,000	\$3,600,355	6
The Father	Drama	97	Trademark Films, Cine@, AG Studios NYC	26 February 2021	\$10,000	n/a	25

PREDICTIONS: 2021 Best Picture Winners

Title	Genres	Runtime	Produced by	Release date	Budget (USD)	US open WE	prev_Oscars
The Trial of the Chicago 7	Drama,History,Thriller	129	Dreamworks Pictures, Amblin Partners, Aperture Media Partners	16 October 2020	\$10,000	n/a	20
Tenet	Action,Sci-Fi,Thriller	150	Warner Bros., Syncopy	3 September 2020	\$205,000,000	\$20,200,000	19
I Care a Lot	Comedy,Crime,Thriller	118	Black Bear Pictures, Crimple Beck	19 February 2021	\$10,000	n/a	11
The Turning	Drama,Horror,Mystery	94	Dreamworks Pictures, Reliance Entertainment, Chislehurst Entertainment Productions	24 January 2020	\$14,000,000	\$6,950,045	3
Like a Boss	Comedy	83	Artists First, Paramount Pictures	10 January 2020	\$29,000,000	\$10,011,272	4
The Last Thing He Wanted	Crime,Drama,Thriller	115	Elevated Films (II), Pimienta, The Fyzz Facility	21 February 2020	\$10,000	n/a	28
Clouds	Drama,Music	121	Mad Chance, Warner Bros., Wayfarer Studios	16 October 2020	\$10,000	n/a	2
Wonder Woman 1984	Action,Adventure,Fantasy	151	Atlas Entertainment, DC Comics, DC Entertainment	25 December 2020	\$10,000	n/a	16
Gekijouban Fate/Stay Night: Heaven's Feel - III. Spring Song	Action,Animation,Drama	120	Aniplex, Type-Moon, Ufotable	18 November 2020	\$10,000	\$200,000	0
The Midnight Sky	Drama,Fantasy,Sci-Fi	118	Anonymous Content, Netflix, Smokehouse Pictures	23 December 2020	\$10,000	n/a	26

So now what?

Some areas for
further exploration

- ◇ Other aspects of diversity in the industry and the Oscars
- ◇ Analysis of results by category (e.g. Best Cinematography)
- ◇ Bayes: probability of nomination given Golden Globe nomination award given nomination, etc.
- ◇ Review and study clusters for archetypes of “Oscar movies”

THANKS!

What questions can I answer?



CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**

Please keep this slide for attribution