POPULARITY ON WIKIPEDIA

orgil batzaya | emily falcon

steven herrera | salvador chavero
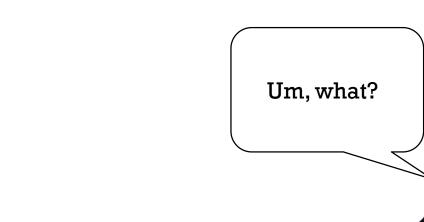
	article_id	full_name	sex	birth_year	city	state	country	continent	latitude	longitude
1	308	Aristotle	Male	-384	Stageira		Greece	Europe	40.33333	23.50000
2	22954	Plato	Male	-427	Athens		Greece	Europe	37.96667	23.71667
3	1095706	Jesus Christ	Male	-4	Judea		Israel	Asia	32.50000	34.90000
4	25664190	Socrates	Male	-469	Athens		Greece	Europe	37.96667	23.71667
5	783	Alexander the Great	Male	-356	Pella		Greece	Europe	40.80000	22.51667
6	18079	Leonardo da Vinci	Male	1452	Vinci		Italy	Europe	43.78333	10.91667
7	5823	Confucius	Male	-551	Qufu		China	Asia	NA	NA
8	15924	Julius Caesar	Male	-100	Rome		Italy	Europe	41.90000	12.50000
9	13633	Homer	Male	-800	Smyrna		Turkey	Europe	38.41861	27.13917
10	23275	Pythagoras	Male	-570	Samos Island		Greece	Europe	37.75000	26.83333
11	1844	Archimedes	Male	-287	Syracuse		Italy	Europe	37.08333	15.28333
12	19577	Moses	Male	-1391	Wadi Tumilat		Egypt	Africa	NA	NA
13	18934	Muhammad	Male	570	Mecca		Saudi Arabia	Asia	21.41667	39.81667
14	1436	Abraham	Male	-1800	Other	NA			33.70000	43.50000
15	27731583	Adolf Hitler	Male	1889	Braunau am Inn		Austria	Europe	48.25833	13.03500
16	33163	Wolfgang Amadeus...	Male	1756	Salzburg		Austria	Europe	47.80000	13.03333
17	5314	Charlemagne	Male	742	Liège		Belgium	Europe	NA	NA
18	32897	William Shakespeare	Male	1564	Stratford-upon-Avon		United Kingdom	Europe	52.19000	-1.71000
19	21019	Michelangelo	Male	1475	Caprese Michelangelo		Italy	Europe	43.63333	11.98333
20	1273	Augustus	Male	-63	Rome		Italy	Europe	41.90000	12.50000
21	69880	Napoleon Bonaparte	Male	1769	Corsica		France	Europe	41.92670	8.73690
22	14627	Isaac Newton	Male	1642	Woolsthorpe-by-Colsterworth		United Kingdom	Europe	52.80986	-0.62877
23	736	Albert Einstein	Male	1879	Ulm		Germany	Europe	48.40000	9.98333
24	5635	Christopher Colum...	Male	1451	Genoa		Italy	Europe	44.41111	8.93278
25	9906294	Johann Sebastian B...	Male	1685	Eisenach		Germany	Europe	50.97611	10.32056
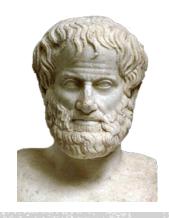
occupation	industry	domain	article_languages	page_views	average_views	historical_popularity_index
Philosopher	Philosophy	Humanities	152	56355172	370758	31.9938
Philosopher	Philosophy	Humanities	142	46812003	329662	31.9888
Religious Figure	Religion	Institutions	214	60299092	281771	31.8981
Philosopher	Philosophy	Humanities	137	40307143	294213	31.6521
Military Personnel	Military	Institutions	138	48358148	350421	31.5840
Inventor	Invention	Science & Technology	174	88931135	511098	31.4644
Philosopher	Philosophy	Humanities	192	22363652	116477	31.3705
Politician	Government	Institutions	128	43088745	336631	31.1161
Writer	Language	Humanities	141	20839405	147797	31.1087
Philosopher	Philosophy	Humanities	114	26168219	229546	31.0691
Mathematician	Math	Science & Technology	129	22022077	170714	30.9907
Religious Figure	Religion	Institutions	115	21880195	190263	30.8632
Religious Figure	Religion	Institutions	150	43026648	286844	30.6500
Religious Figure	Religion	Institutions	105	17824676	169759	30.6047
Politician	Government	Institutions	169	128368798	759579	30.5778
Composer	Music	Arts	177	51363334	290188	30.5146
Politician	Government	Institutions	116	26626252	229537	30.4837
Writer	Language	Humanities	163	74882297	459401	30.4398
Painter	Fine Arts	Arts	158	39992803	253119	30.4383
Politician	Government	Institutions	117	20307663	173570	30.4195
Politician	Government	Institutions	145	57559307	396961	30.3263
Physicist	Natural Sciences	Science & Technology	191	59909049	313660	30.2889
Physicist	Natural Sciences	Science & Technology	166	89771090	540790	30.2106
Explorer	Explorers	Exploration	153	58442090	381974	30.1753
Composer	Music	Arts	144	29753654	206623	30.1672

INTRODUCTION TO OUR DATA: PANTHEON
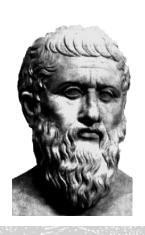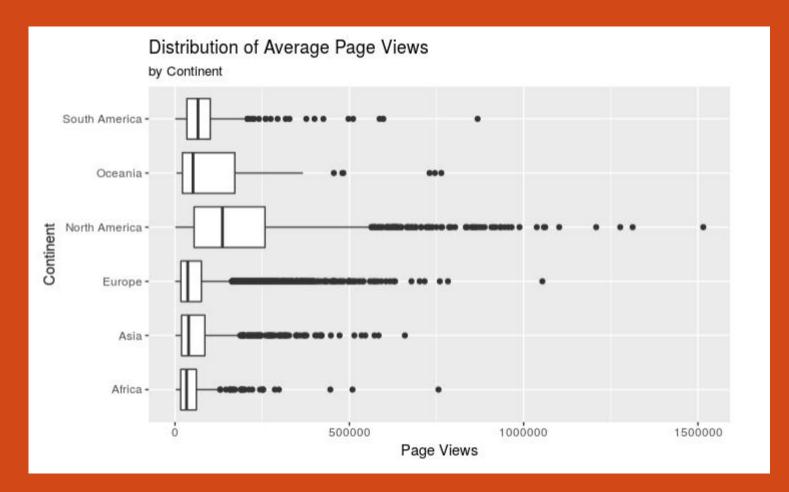
WHAT VARIABLES INFLUENCE THE HISTORICAL POPULARITY INDEX?

Our analysis through linear regression

Historical Popularity Index Around the World

Distribution of Average Page Views by Continent

PAGE VIEWS

Manny Pacquiao
(Asia)

Chris Hemsworth
(Oceania)

Cristiano Ronaldo
(Europe)

Kim Kardashian
(North America)

Lionel Messi
(South America)
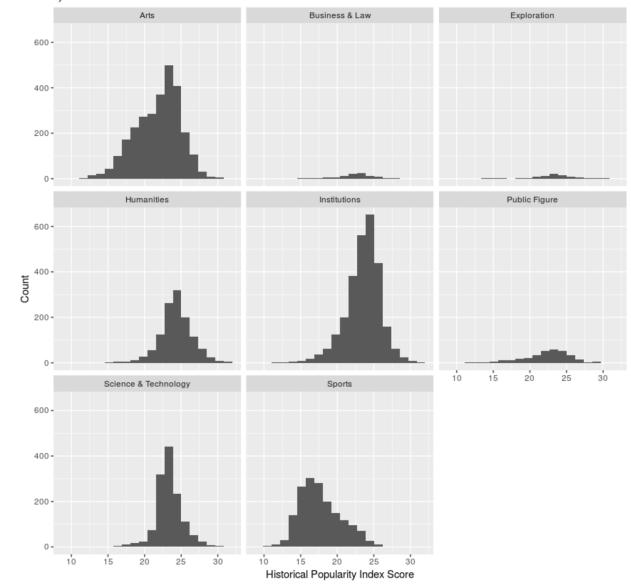
Freddie Mercury
(Africa)

OUTLIERS

SECTION 1:

DOMAIN

Distribution of Historical Popularity Index Scores by Domain

DOMAIN

- Levels:
  - Arts
  - Institutions
  - Sports
  - Science & Technology
  - Humanities
  - Public Figure
  - Business & Law
  - Exploration

Historical Popularity Index

$$= 21.839(\text{intercept})$$
$$+ 0.502(\text{Business \& Law})$$
$$+ 1.710(\text{Exploration})$$
$$+ 2.448(\text{Humanities})$$
$$+ 1.716(\text{Institutions})$$
$$+ 0.532(\text{Public Figure})$$
$$+ 1.534(\text{Science \& Technology})$$
$$+ 4.112(\text{Sports})$$

DOMAIN

SECTION 2:

SEX

Not that kind…

SEX

Historical Popularity Index (-3500 – 2005)

=

21.839(intercept)

+

1.5535(sexMale)

Historical Popularity Index (post 1920)

=
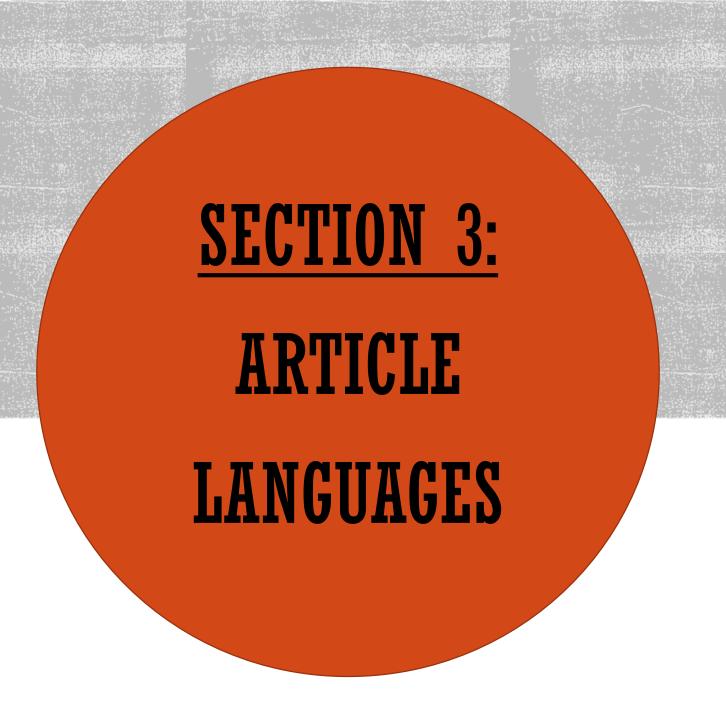
21.839(intercept)

+

0.5687(sexMale)

- 1,427 women (3500 BCE to 2005)
- 8,852 men (3500 BCE to 2005)

- 1,427 women

- 1,053 women (1920 to 2005)
- 4,309 men (1920 BCE to 2005)

SECTION 3:

ARTICLE

LANGUAGES

Historical Popularity Index
=
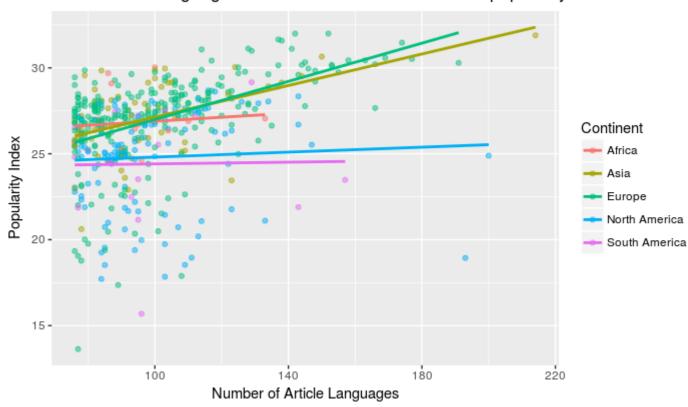21.839(intercept)
+
0.089(article_languages)

ARTICLE LANGUAGES

Effect of # of languages the article is translated into on popularity index

ARTICLE LANGUAGES

FULL LINEAR MODEL ANALYSIS

Historical Popularity Index

$$
=
$$

20.79(intercept)

$+$

1.431(sexMale)

$+$

0.335(domainBusiness & Law)

$+$

0.291(domainExploration)

$+$

1.099(domainHumanities)

$+$

0.3941(domainInstitutions)

$+$

0.7533(domainPublic Figure)

$+$

0.7814(domainScience & Technology)

$-$

4.204(domainSports)

$+$

0.0675(article_languages)

$+$

0.290(continentAsia)

$+$

1.224(continentEurope)

$+$

0.4113(continent North America)

$-$

0.204(continentOceania)

$+$

1.155(continentSouth America)

$-$

0.00182(birth_year)

FULL/SELECTED MODEL

BASED ON THE MODEL, A FIGURE THAT MEETS THESE CHARACTERISTICS SHOULD HAVE THE HIGHEST PREDICTED POPULARITY SCORE

- Male
- From Europe
- Works in humanities
- Live before the year 0

WHO ARE THEY?

```{r perfect_index_score}
panth_mod %>%
  filter(continent == "Europe",
         birth_year < 0,
         sex == "Male",
         domain == "Humanities") %>%
    select(rank,full_name,historical_popularity_index, occupation)
```

rank <int>	full_name <chr>	historical_popularity_index <dbl>	occupation <chr>
1	Aristotle	31.9938	Philosopher
2	Plato	31.9888	Philosopher
4	Socrates	31.6521	Philosopher
8	Homer	31.1087	Writer
9	Pythagoras	31.0691	Philosopher
30	Virgil	29.8811	Writer
35	Sophocles	29.7421	Writer
46	Epicurus	29.4762	Philosopher
53	Ovid	29.2723	Writer
66	Aeschylus	29.0996	Writer

CONCLUDING REMARKS

Weaknesses

- Dataframe is a population
- Variables not tailored for certain analyses

Future Goals

- Shiny App visual
- Timeline visualization/analysis
- Creating world map
- Web-scraping