

MA615 Tidyverse Problem Set

Shixin Liang

10/3/2019

Problem 1

How many continents are included in the data set?

```
## [1] 5
```

How many countrys are included?

```
## [1] 142
```

How many countries per continent?

```
## [1] 52
```

```
## [1] 25
```

```
## [1] 33
```

```
## [1] 30
```

```
## [1] 2
```

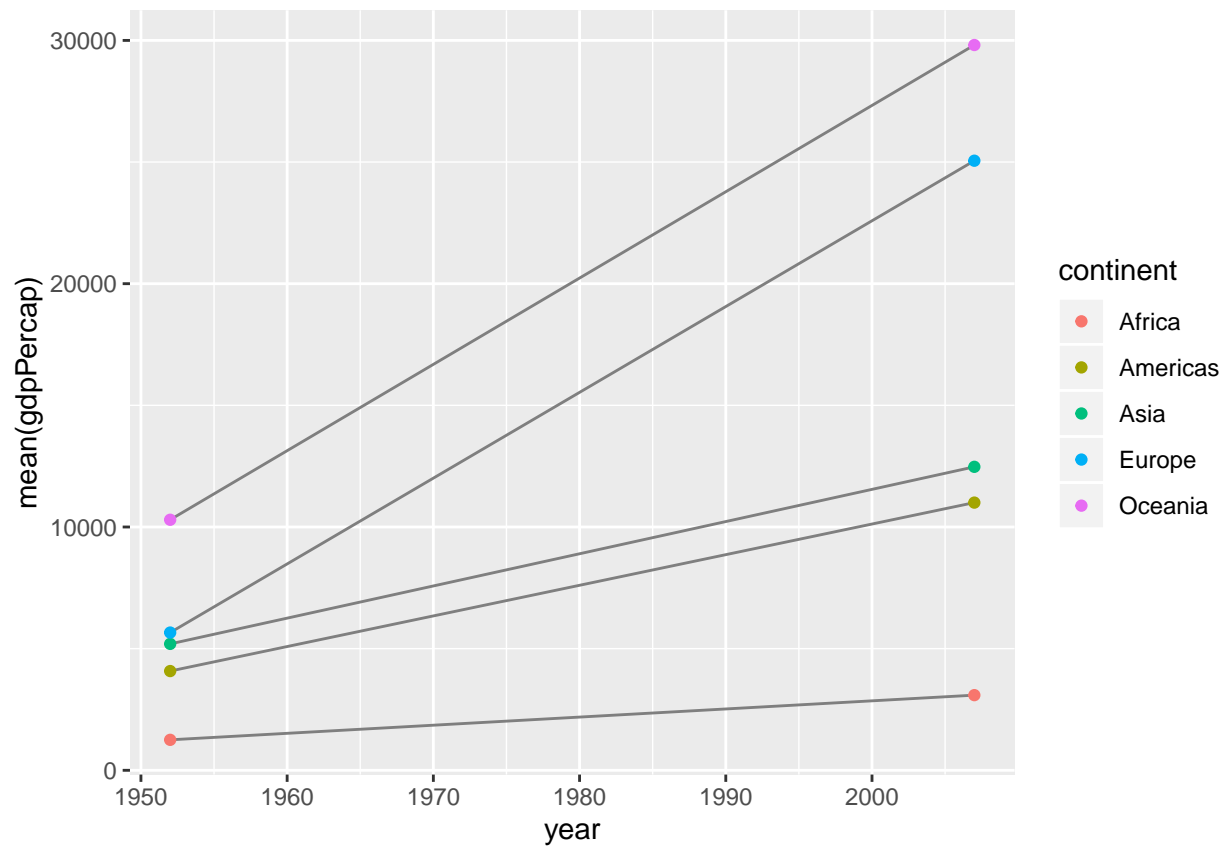
A report showing the continents in the dataset, total population per continent, and GDP per capita.

continent	n	total_GDP\$n
Africa	6187585961	1368902.9
Americas	7351438499	2140833.1
Asia	30507333901	3129251.6
Europe	6181115304	5209011.2
Oceania	212992136	446918.6

Produce a well-labeled table that summarizes GDP per capita for the countries in each continent, contrasting the years 1952 and 2007.

year	continent	mean(gdpPercap)
1952	Africa	1252.572
1952	Americas	4079.063
1952	Asia	5195.484
1952	Europe	5661.057
1952	Oceania	10298.086
2007	Africa	3089.033
2007	Americas	11003.032
2007	Asia	12473.027
2007	Europe	25054.482
2007	Oceania	29810.188

Product a plot that summarizes the same data as the table.



Which countries in the dataset have had periods of negative population growth?
Here are the countries:

country	continent	year	lifeExp	pop	gdpPercap	popular_change
Afghanistan	Asia	1982	39.854	12881816	978.0114	-1998556
Albania	Europe	1952	55.230	1282697	1601.0561	-30607226
Angola	Africa	1952	30.015	4232095	3520.6103	-29101121
Australia	Oceania	1952	69.120	8691212	10039.5956	-31610715
Austria	Europe	1952	66.800	6927772	6137.0765	-13506404
Bahrain	Asia	1952	50.939	120447	9867.0848	-8079336
Belgium	Europe	1952	68.000	8730405	8343.1051	-141717934
Benin	Africa	1952	38.223	1738315	1062.7522	-8653911
Bolivia	Americas	1952	40.414	2883315	2677.3263	-5194999
Bosnia and Herzegovina	Europe	1952	53.820	2791000	973.5332	-6328152
Bosnia and Herzegovina	Europe	1992	72.178	4256013	2546.7814	-82964
Bosnia and Herzegovina	Europe	1997	73.244	3607000	4766.3559	-649013
Botswana	Africa	1952	47.622	442308	851.2411	-4109890
Bulgaria	Europe	1952	59.600	7274900	2444.2866	-182735747
Bulgaria	Europe	1992	71.190	8658506	6302.6234	-313452
Bulgaria	Europe	1997	70.320	8066057	5970.3888	-592449
Bulgaria	Europe	2002	72.140	7661799	7696.7777	-404258
Bulgaria	Europe	2007	73.005	7322858	10680.7928	-338941
Burkina Faso	Africa	1952	31.975	4469979	543.2552	-2852879
Burundi	Africa	1952	39.031	2445618	339.2965	-11880585
Cambodia	Asia	1952	39.417	4693836	368.4693	-3696669
Cambodia	Asia	1977	31.220	6978607	524.9722	-471999
Cameroon	Africa	1952	38.523	5009067	1172.6677	-9122791
Canada	Americas	1952	68.750	14785584	11367.1611	-2910709
Central African Republic	Africa	1952	35.463	1291695	1071.3107	-32098446
Chad	Africa	1952	38.092	2682462	1178.6659	-1686576
Chile	Americas	1952	54.745	6377619	3939.9788	-3861188
Colombia	Americas	1952	50.643	12350771	2144.1151	-1306332325
Comoros	Africa	1952	40.715	153936	1102.9909	-44073614
Congo, Rep.	Africa	1952	42.111	854885	2125.6214	-63751874
Costa Rica	Americas	1952	57.206	926317	2627.0095	-2874293
Cote d'Ivoire	Africa	1952	40.477	2977019	1388.5947	-1156865
Croatia	Europe	1952	61.210	3882229	3119.2365	-14131180
Croatia	Europe	1997	73.680	4444595	9875.6045	-49418
Czech Republic	Europe	1952	66.870	9125183	6876.1403	-2291804
Czech Republic	Europe	1997	74.010	10300707	16048.5142	-14995
Czech Republic	Europe	2002	75.510	10256295	17596.2102	-44412
Czech Republic	Europe	2007	76.486	10228744	22833.3085	-27551
Denmark	Europe	1952	70.780	4334000	9692.3852	-5894744
Djibouti	Africa	1952	34.812	63149	2669.5295	-5404971
Ecuador	Americas	1952	48.357	3548753	3522.1107	-5770869
El Salvador	Americas	1952	45.262	2042865	3048.3029	-78221678
Equatorial Guinea	Africa	1952	34.482	216964	375.6431	-6722724
Equatorial Guinea	Africa	1977	42.024	192675	958.5668	-84928
Finland	Europe	1952	66.550	4090500	6424.5191	-72421387
Gabon	Africa	1952	37.003	420702	4293.4765	-60663214
Gambia	Africa	1952	30.000	284320	485.2307	-1170547
Germany	Europe	1977	72.500	78160773	20512.9212	-556315
Germany	Europe	1987	74.847	77718298	24639.1857	-616968
Ghana	Africa	1952	43.149	5581001	911.2989	-76819995
Greece	Europe	1952	65.860	7733250	3530.6901	-15140088
Guatemala	Americas	1952	42.023	3146381	2428.2378	-7559909
Guinea	Africa	1952	33.609	2664249	510.1965	-9908679
Guinea-Bissau	Africa	1952	32.590	580653	299.8503	-9367161
Guinea-Bissau	Africa	1967	35.492	601287	715.5806	-26533
Honduras	Americas	1952	41.912	1517453	2194.9262	-6985361
Hong Kong, China	Asia	1952	60.960	2125900	3054.4212	-5357865
Hong Kong, China	Asia	1987	69.538	19812748	12883.4388	-92785

##	[1]	Afghanistan	Albania
##	[3]	Angola	Australia
##	[5]	Austria	Bahrain
##	[7]	Belgium	Benin
##	[9]	Bolivia	Bosnia and Herzegovina
##	[11]	Botswana	Bulgaria
##	[13]	Burkina Faso	Burundi
##	[15]	Cambodia	Cameroon
##	[17]	Canada	Central African Republic
##	[19]	Chad	Chile
##	[21]	Colombia	Comoros
##	[23]	Congo, Rep.	Costa Rica
##	[25]	Cote d'Ivoire	Croatia
##	[27]	Czech Republic	Denmark
##	[29]	Djibouti	Ecuador
##	[31]	El Salvador	Equatorial Guinea
##	[33]	Finland	Gabon
##	[35]	Gambia	Germany
##	[37]	Ghana	Greece
##	[39]	Guatemala	Guinea
##	[41]	Guinea-Bissau	Honduras
##	[43]	Hong Kong, China	Hungary
##	[45]	Iceland	Indonesia
##	[47]	Iran	Iraq
##	[49]	Ireland	Israel
##	[51]	Jamaica	Jordan
##	[53]	Korea, Dem. Rep.	Korea, Rep.
##	[55]	Kuwait	Lebanon
##	[57]	Lesotho	Liberia
##	[59]	Libya	Madagascar
##	[61]	Malawi	Malaysia
##	[63]	Mali	Mauritania
##	[65]	Mauritius	Mongolia
##	[67]	Montenegro	Mozambique
##	[69]	Namibia	Netherlands
##	[71]	New Zealand	Nicaragua
##	[73]	Niger	Norway
##	[75]	Oman	Panama
##	[77]	Paraguay	Philippines
##	[79]	Poland	Portugal
##	[81]	Puerto Rico	Reunion
##	[83]	Romania	Rwanda
##	[85]	Sao Tome and Principe	Senegal
##	[87]	Serbia	Sierra Leone
##	[89]	Singapore	Slovak Republic
##	[91]	Slovenia	Somalia
##	[93]	South Africa	Spain
##	[95]	Sri Lanka	Sudan
##	[97]	Swaziland	Switzerland
##	[99]	Syria	Taiwan
##	[101]	Tanzania	Thailand
##	[103]	Togo	Trinidad and Tobago
##	[105]	Uganda	Uruguay
##	[107]	West Bank and Gaza	Zambia

```
## [109] Zimbabwe
## 142 Levels: Afghanistan Albania Algeria Angola Argentina ... Zimbabwe
```

Which countries in the dataset have had the highest rate of growth in per capita GDP?

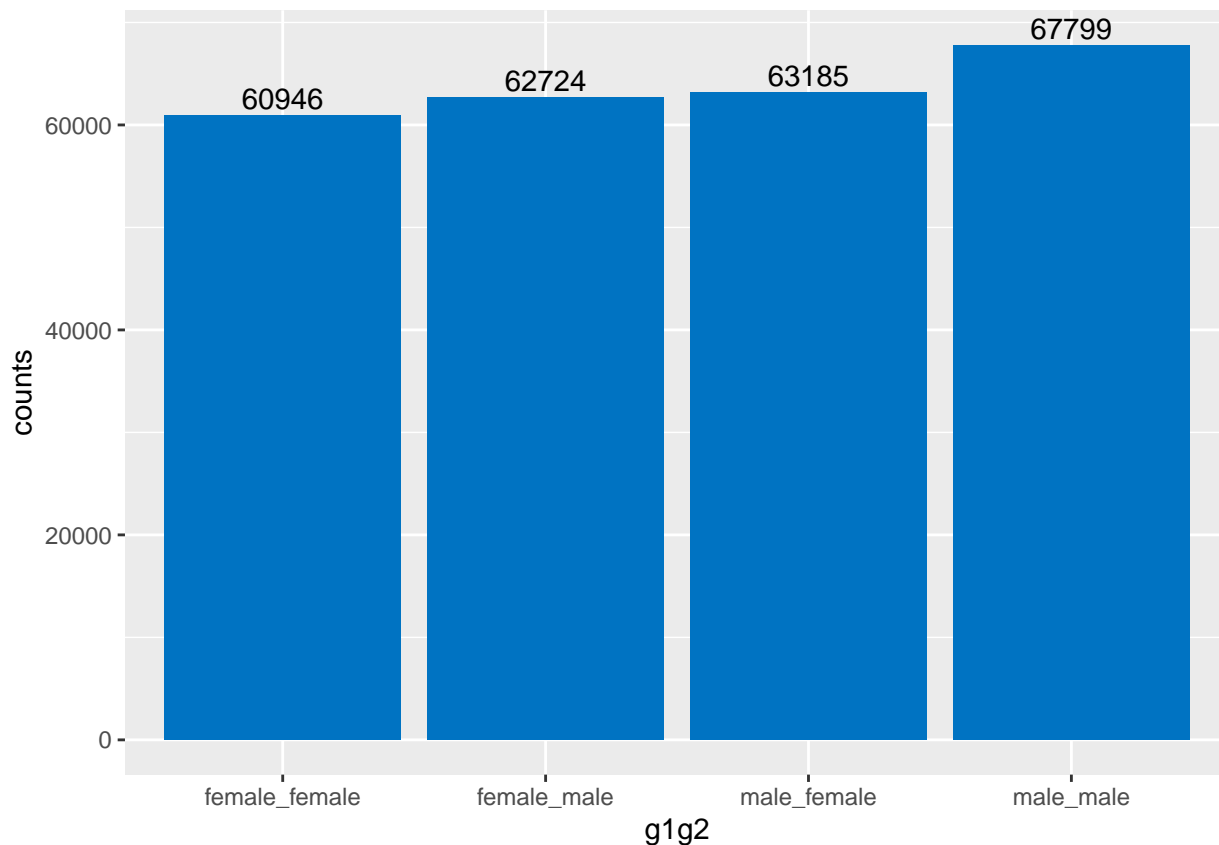
country	continent	year	lifeExp	pop	gdpPercap	popular_change	GDP_change
Kuwait	Asia	1952	55.565	160000	108382.4	-48884790	85034.21

Kuwait had the highest rate.

Problem 2

Product a plot the contracts the frequency of these four gendercombinations.

```
## Loading required package: car
## Loading required package: carData
##
## Attaching package: 'car'
## The following object is masked from 'package:dplyr':
##
##      recode
## The following object is masked from 'package:purrr':
##
##      some
## Loading required package: lmtest
## Loading required package: zoo
##
## Attaching package: 'zoo'
## The following objects are masked from 'package:base':
##
##      as.Date, as.Date.numeric
## Loading required package: sandwich
## Loading required package: survival
## [1] male_female   female_male   female_female male_male
## Levels: female_female female_male male_female male_male
```



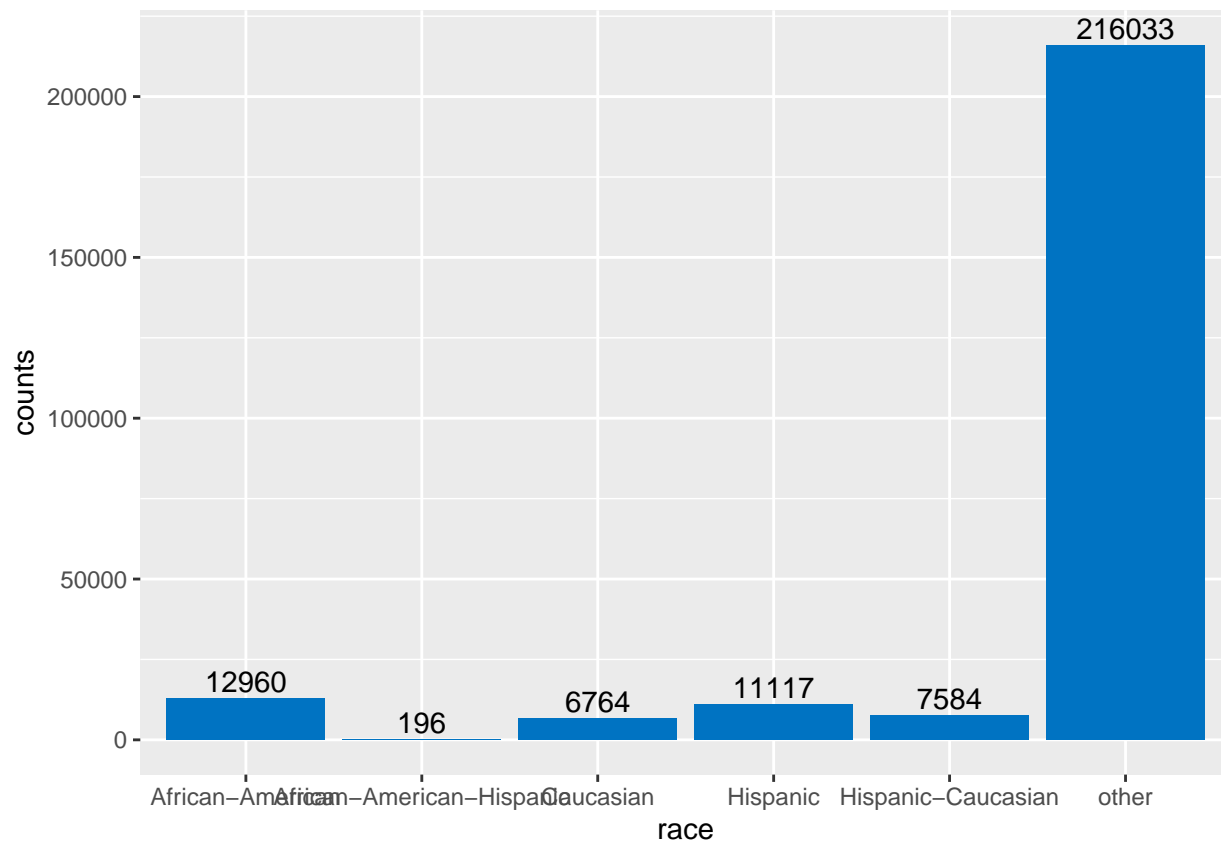
Are the frequencies different for women in their 20s and women who are older than 29?

g1g2	counts	age
female_female	22183	20s
female_male	22508	20s
male_female	22653	20s
male_male	24505	20s
female_female	38763	30s
female_male	40216	30s
male_female	40532	30s
male_male	43294	30s

The frequencies for women who are older than 29 are higher.

Produce a plot that contrasts the frequency of having more than two children by race and ethnicity.

```
## [1] other                African-American
## [3] Caucasian              Hispanic
## [5] African-American-Hispanic Hispanic-Caucasian
## 6 Levels: African-American African-American-Hispanic ... other
```



Problem 3

How many times does the letter “e” occur in mtcars rownames?

```
## [1] "Hornet 4 Drive"      "Hornet Sportabout"  "Duster 360"
## [4] "Merc 240D"           "Merc 230"           "Merc 280"
## [7] "Merc 280C"           "Merc 450SE"         "Merc 450SL"
## [10] "Merc 450SLC"         "Cadillac Fleetwood" "Lincoln Continental"
## [13] "Chrysler Imperial"  "Dodge Challenger"   "AMC Javelin"
## [16] "Pontiac Firebird"    "Porsche 914-2"       "Ford Pantera L"
## [19] "Ferrari Dino"        "Maserati Bora"
## [1] 25
```

How many cars in mtcars have the brand Merc?

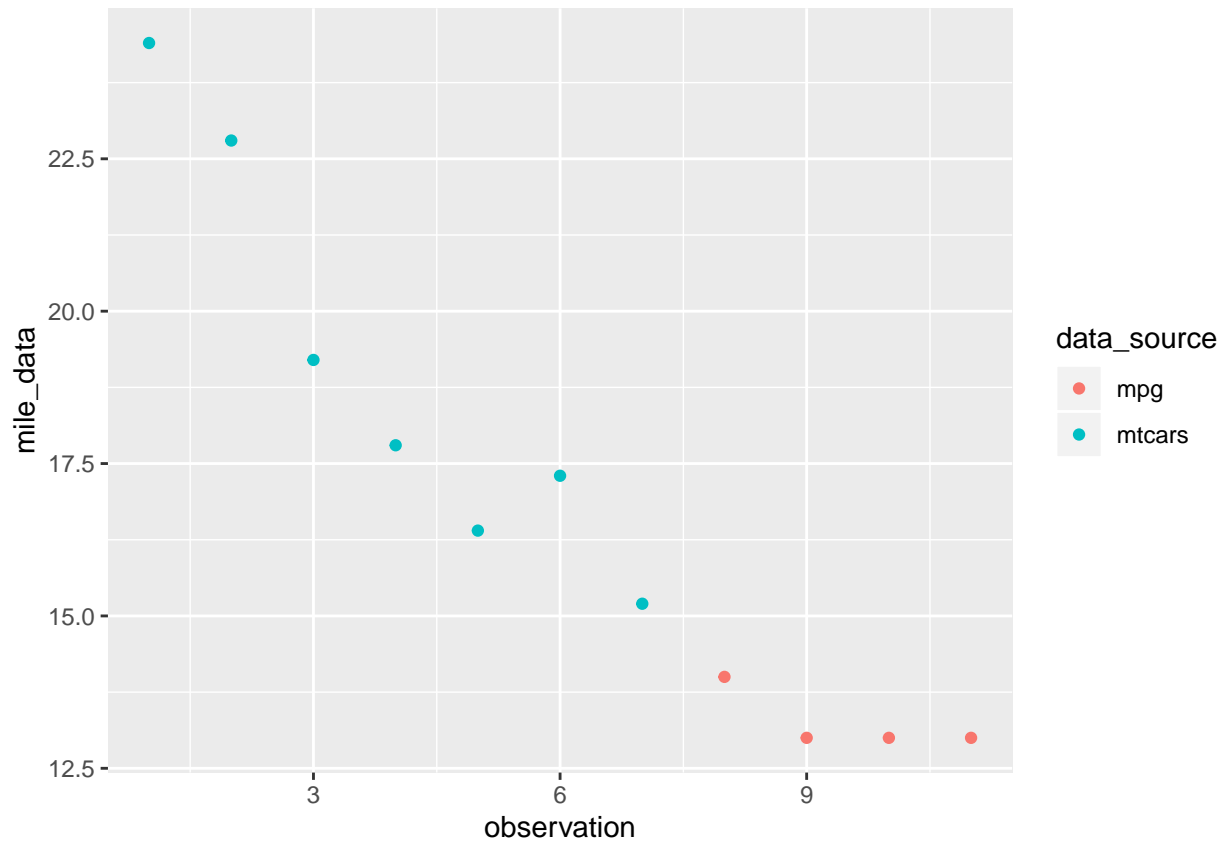
```
## [1] 7
```

How many cars in mpg have the brand(“manufacturer” in mpg) Merc?

```
## [1] 4
```

Contrast the mileage data for Merc cars as reported in mtcars and mpg. Use tables, plots, and a short explanation.

```
## [1] 7
```



The mileage data for Merc cars from mpg is smaller than those from mtcars. It is because the mileage data from mpg are mileage data in city, not in highway.

Problem 4

Produce a tibble that displays the five most popular boy names and girl names in the years 1880,1920, 1960, 2000.

```
## [1] "John"      "William" "James"   "Charles" "George"
## [1] "Mary"      "Anna"    "Emma"    "Elizabeth" "Minnie"
## [1] "John"      "William" "Robert"  "James"   "Charles"
## [1] "Mary"      "Dorothy" "Helen"   "Margaret" "Ruth"
## [1] "David"     "Michael" "James"   "John"    "Robert"
## [1] "Mary"     "Susan" "Linda" "Karen" "Donna"
## [1] "Jacob"     "Michael" "Matthew" "Joshua"  "Christopher"
## [1] "Emily"     "Hannah" "Madison" "Ashley"  "Sarah"
```

What names overlap boys and girls?

```
## # A tibble: 10,663 x 1
##   name
##   <chr>
## 1 John
## 2 William
## 3 James
```



```
## 4 Charles
## 5 George
## 6 Frank
## 7 Joseph
## 8 Thomas
## 9 Henry
## 10 Robert
## # ... with 10,653 more rows
```

Produce a chart that shows the relative frequency of the names “Donald”, “Hilary”, “Hillary”, “Joe”, “Barrack”, over the years 1880 through 2017.

