

Denzel

27 marks

The actor [Denzel Washington](#) has appeared in many movies. In some he appears with facial hair, or with a hat, or wearing glasses, or some combination of these.

Download the `denzel.R` file and load it directly into R using `source("denzel.R")` (you may have to give the full path to the file if this doesn't work, or just open the file in R and execute its contents).

This will provide data on twenty-one movies in which Mr. Washington has appeared and what sort of choice of accessory (i.e. hat, facial hair, glasses) he was sporting in each movie. The data is given in two forms, one as a data frame called `denzel` whose `row.names(denzel)` will give the movie titles. Its three variates are binary, each indicating whether Mr. Washington sported one or more of the three possibilities of facial hair, glasses, or a hat. The other contains the same information (absent the movie titles) but as a table of counts called `denzelTable`.

- a. **Venn diagrams.** Install the package `venneuler` from the CRAN repository. (You will need `rJava` and hence `Java` installed. If this doesn't work, use the windows machines run by MFCF on the third floor of the MC building.)
 - i. **(4 marks)** Using `venneuler(...)`, draw a Venn diagram for the three sets "facial hair", "glasses", and "hat". **Hint:** you will need to determine an appropriate vector of `movies` and a matching vector of `choices` so that you can create an appropriate data frame `venndenzel` and produce a plot as

```
venn <- venneuler(venndenzel)
plot(venn)
```

In constructing `movies` and `choices` it might be best to construct three separate pairs of vectors first, one for each choice and then concatenate them together. For example,

```
facialhair <- denzel[, "facialhair"] == "yes"
facialhairMovies <- row.names(denzel)[facialhair]
facialhairChoices <- rep("facial hair", length(facialhairMovies))
# et cetera
movies <- c(facialhairMovies, hatMovies, glassesMovies)
choices <- c(facialhairChoices, hatChoices, glassesChoices)
```

- ii. **(2 marks)** What can you conclude from this diagram about the relationship between these choices?
 - iii. **(2 marks)** By looking at the residuals of the output from `venneuler(...)`, which set of choices is best represented by area in this diagram? Which is worst?
- b. **Barplots** Now using the data contained in `denzelTable`,
 - i. **(3 marks)** Construct and print each of the marginal tables: `facialhair × hat`, `facialhair × glasses`, and `hat × glasses`.
 - ii. **(3 marks)** Produce a barplot from each of the above tables showing `facialhair` by `hat`, `facialhair` by `glasses`, and `hat` by `glasses`, respectively. Have a legend attached to each.
 - iii. **(3 marks)** What conclusions can you draw about the relationship between the choices from these barplot displays? How do these displays compare to the Venn diagrams above in the information they provide?
- c. **Eikosograms.** You will need to first install the package `eikosograms` (from CRAN). For this question, you will be working with the data frame `denzel`.
 - i. **(4 marks)** Produce the eikosogram for each the following combinations
 - Facial hair conditional only on hat or not.
 - Facial hair conditional only on glasses or not.
 - Hat conditional on glasses or not.What conclusions do you draw about these various conditional distributions?
 - ii. **(6 marks)** Produce the eikosogram for each the following combinations
 - Facial hair conditional on all possible choices of glasses and hat
 - Glasses conditional on all possible choices of facial hair and hat
 - Hat conditional on all possible choices of glasses and facial hairWhat conclusions do you draw about each of these various conditional distributions?