Dear all,

Here are detailed instruction for Project-Report #2, due by end of today. Please submit your report and codes into the folder of



Step1: With the given “morph\_2008\_nonCommercial.csv”, which is originally provided,

1. What is the distinct # of Males?

11459

1. What is the distinct # of Females?

2158

1. What is the distinct # of subjects? What is your conclusion based on (a) to (c)?

13617, this is not equal to a+b which is 13618

1. Please fill out the following table:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Black | White | Asian | Hispanic | Other | Total |
| Male | 8832 | 2060 | 47 | 507 | 13 | 11459 |
| Female | 1492 | 632 | 3 | 27 | 5 | 2158 |
| Total | 10323 | 2692 | 50 | 534 | 18 | 13617 |

Step2: With the “morphII\_cleaned\_v2.csv” from 2017 REU program, answer the same questions from above:

1. What is the distinct # of Males?

11458

1. What is the distinct # of Females?

2159

1. What is the distinct # of subjects? What is your conclusion based on (a) to (c)?

13617, a+b is equal to c!

1. Please fill out the following table:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Black | White | Asian | Hispanic | Other | Total |
| Male | 8829 | 2056 | 47 | 507 | 19 | 11458 |
| Female | 1491 | 628 | 4 | 28 | 8 | 2159 |
| Total | 10320 | 2684 | 51 | 535 | 27 | 13617 |

Step3: With the given “morphII\_cleaned\_v2.csv”, answer the following questions:

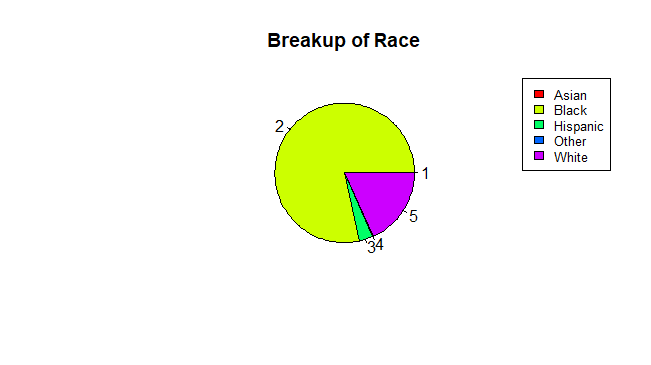
1. Find the number of additional images per subject:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5+ | Total |
| Male | 2350 | 3606 | 1975 | 1135 | 2020 | 11086 |
| Female | 478 | 712 | 352 | 172 | 360 | 2074 |
| Total | 2828 | 4318 | 2327 | 1307 | 2390 | 13160 |

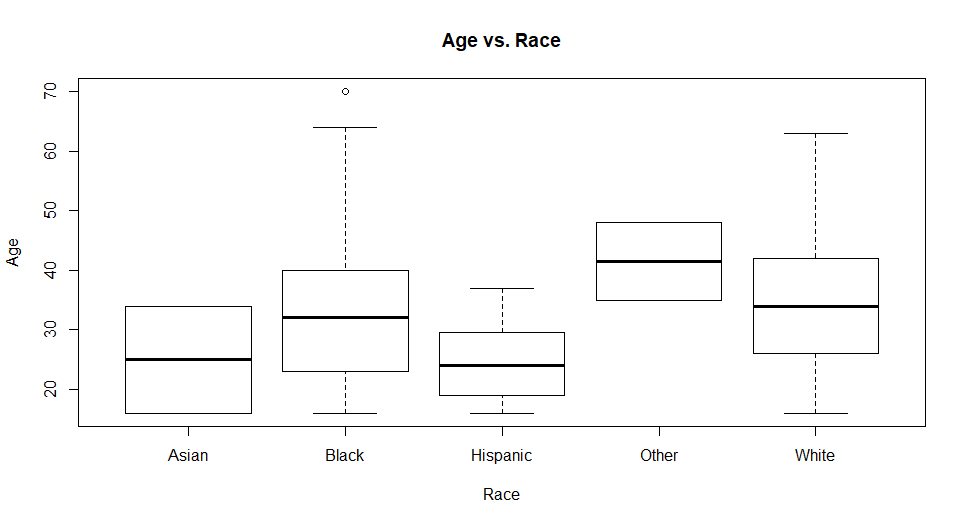
1. Find the number of Facial images by gender and Decade-of-Life

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | <20 | 20-29 | 30-39 | 40-48 | 50+ | Total |
| Male | 1968 | 3387 | 3047 | 2165 | 891 | 11458 |
| Female | 294 | 607 | 686 | 451 | 121 | 2159 |
| Total | 2262 | 3994 | 3733 | 2616 | 1012 | 13617 |

Step4: With the “morphII\_cleaned\_v2.csv”, and “MorphII\_BIF\_s7-37\_g0.1\_max\_partial.csv”, merge both files together to get the information of Race. ***Provide numerical and graphical summary to Race, which is similar to what you completed from Project 1***.



*Figure 1: Breakup of Race of the Morph-II Sample Data*



*Figure 2: A Boxplot showing the Distributions of Ages based on Race for the Morph-II Sample Data*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Five Point Summary for Age vs. Race Data | | | | | |
| **Race** | **Asian** | **Black** | **Hispanic** | **Other** | **White** |
| **Min** | *16* | *16* | *16* | *35* | *16* |
| **Q1** | *20.5* | *23* | *19* | *38.25* | *26* |
| **Median** | *25* | *32* | *24* | *41.50* | *34* |
| **Mean** | *25* | *32.23* | *24.69* | *41.50* | *34.52* |
| **Q3** | *29.5* | *40* | *29.25* | *44.75* | *42* |
| **Max** | *34* | *70* | *37* | *48* | *63* |

*Table 1: The Five Point Summary for the Distributions of Age based on Race*