Name Rebecca Tsekanovskiy

RIN 662038623

You need to fill out the two tables below and create histograms of the length and area of each image below the tables to receive credit. Be sure to add figure captions and take-home messages. Graphs should be made in R / ggplot with clear labels etc.

Fill out the below table for each image you analyze. Try to measure at least 10 cells within an image when possible. You may do more if time permits.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Image ID | Subject | Units | Scale Bar? (Y/N) | Scale Conversion | Metrics to measure |
| 36659 | Nucleolus | UM | Y | 0.25 | Length and Area |
| 19129 | Plant Cell Wall | um | y | .25 | Length and Area |
| 36630 | Protozoa | um | y | .25 | Length and Area |

**NOTE: I was told by different TA’s that my data looked right but it was already after I had done all of the scales with the conversion at .25 and I was told to just leave it alone.**

We just measured many cells in each image. It is time to calculate summary statistics to characterize the size of cells in each image. Record your summarized measurements for length **AND** area below for each image below. Add more rows to the table as needed. You should be able to import data into R and then do calculations.

**HINTS:**

* You should have a data set for each cell image.
* I recommend reading each file into R separately.
* You should then filter data in R for length or area measurement. This is confusing because you get values for both metrics for every row. **Measurements of 0 length must be area.**  So, you can filter for area when the length column is equal to 0. For the length measurements, you cold filter by length >0.
* Now calculate your summary statistics and fill in the below table.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Image ID | Metric **(length or area)** you should have two rows for each image  length | Units | Mean | Median | Standard Deviation | Minimum | Maximum |
| 19129 | Area | um | .0610 | .0550 | .03351434 | .0100 | .1100 |
| 19129 | Length | um | .3161 | .2820 | 0.1851213 | .1410 | .7600 |
| 36630 | Area | um | .05760 | .03950 | 0.05479497 | .01100 | .18600 |
| 36630 | Length | um | .2660 | .2360 | . 0.1241155 | .1490 | .5860 |
| 36659 | Area | um | .0201 | .0090 | .03439784 | .0020 | .1170 |
| 36659 | Length | um | .1142 | .1090 | 0.02637676 | .0770 | .1660 |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |

Copy and Paste graphs below. You want to make a histogram for each image one for the length and one for the area (two total for each image). Histograms give you an idea of the distribution of the data set. The code is similar to making a box plot, but this time use geom\_histogram(). In aes() you should make x = to your variable and you should not have a y variable.  See this resource to customize your histogram. <http://www.sthda.com/english/wiki/ggplot2-histogram-plot-quick-start-guide-r-software-and-data-visualization> Please include a figure caption and bullet point take away for each graph. Compare the size of each cell (length and area).

A graph of a number of different sizes

Description automatically generated with medium confidence

Figure Caption: This is a histogram graph of image 19129 demonstrating the area.

* Takeaway: The area goes all the way up to .12 and the highest point is about .05 as it has a frequency of about 3.0. We can see a cluster of cells that have the same area.

A graph of a number of objects

Description automatically generated

Figure Caption: This is a histogram of image 19129 demonstrating the length.

* Takeaway: The length goes all the way up to .8 with the heighest frequency at 4.0. We can see a cluster of cells that have the same length until about .4.

A graph of a number of objects

Description automatically generated

Figure Caption: This is a histogram graph of image 36630 demonstrating the area.

* Takeaway: The area value goes up to around .20 with most of the values clustered around .005. The highest frequency goes up to 5.0. We can see a cluster of cells that have the same area around .05. There are less cells that have an area around .15 and .20.

A graph of a number of objects

Description automatically generated

Figure Caption: This is a histogram of image 36630 demonstrating the length.

* Takeaway: The length value goes all the way up to .6. The highest frequency is at a little over 6 with it being at .20. We can see a cluster of cells that have the same length around .2. All other cells are significantly smaller.

A graph of a number

Description automatically generated

Figure Caption: This is a histogram graph of image 36659 demonstrating the area.

* Takeaway: The area goes all the way to about .120 with the highest freuqency being a little over 6. We can see a cluster of cells that have that very little area with the rest being significantly smaller.

A graph of a number of different sizes

Description automatically generated with medium confidence

Figure Caption: This is a histogram of image 36659 demonstrating the length.

* Takeaway: The length ranges to .150 with the highest freuqnecy being at 3. We can see a cluster of cells that have the same length of about .100 up until .125.