Name: Nicky Lin

RIN: 662041711

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 | 10 | Length and Area |
| 36630 | Protozoa | UM | Y | 0.5 | Length and Area |

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 for 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 | Units | Mean | Median | Standard Deviation | Minimum | Maximum |
| 36659 | area | um^2 | 528.7 | 444.1 | 301.245 | 255.5 | 1060.6 |
| length | um | 24.30 | 22.77 | 6.607286 | 17.14 | 37.49 |
| 19129 | area | um^2 | 77.66 | 71.47 | 26.39153 | 49.00 | 127.48 |
| length | um | 16.104 | 13.738 | 6.144858 | 8.034 | 25.072 |
| 36630 | area | um^2 | 0.2358 | 0.2320 | 0.03825877 | 0.1810 | 0.3000 |
| length | um | 0.5737 | 0.5620 | 0.07129758 | 0.4590 | 0.7290 |

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.

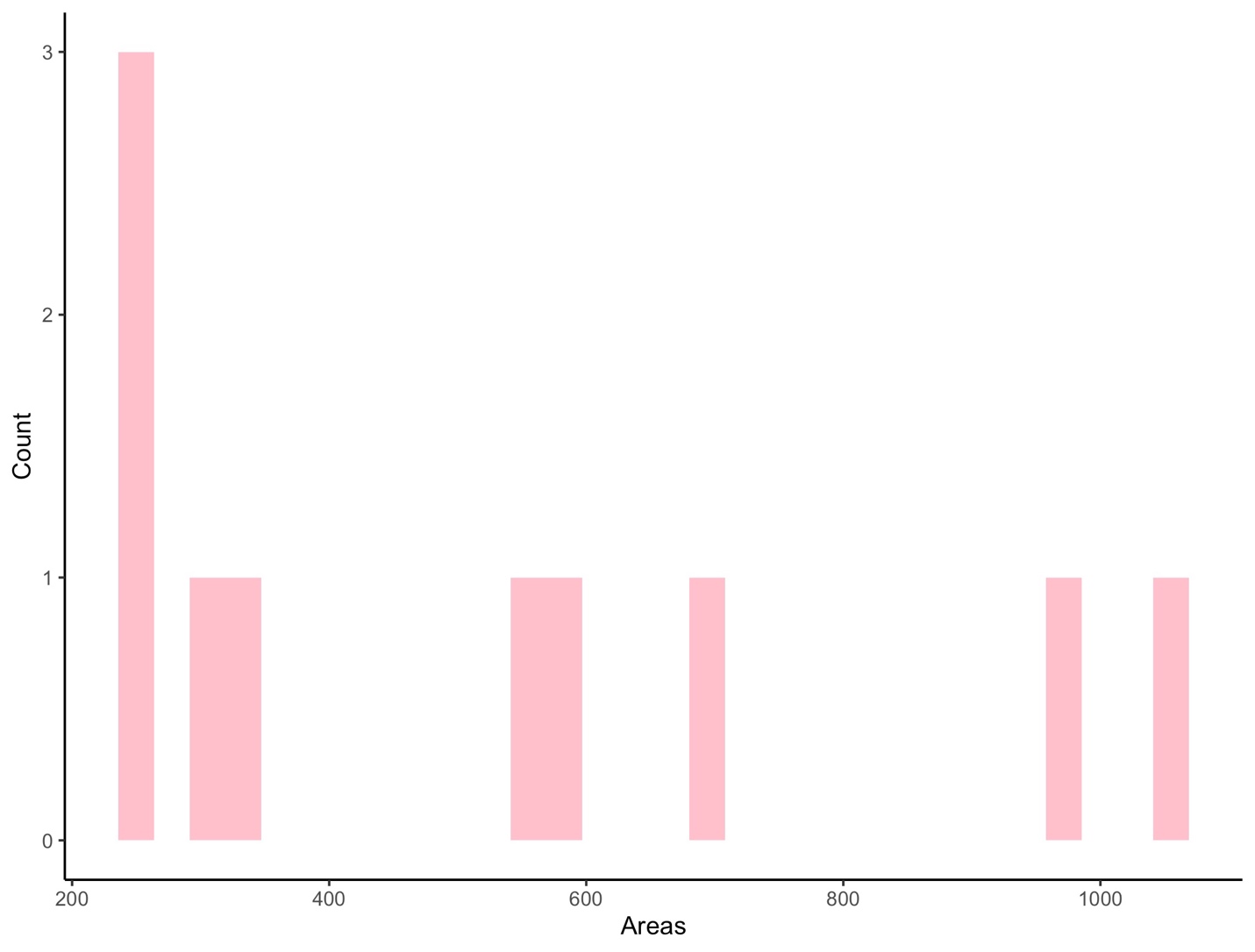


Figure 1. is a histogram of cell areas in the 36659.jpg cell image. 10 cells were randomly chosen and measured in um^2.

Takeaway

* There are relatively more cells that have areas between 200 and 300 um^2. This indicates that small cells are more common.

Chart, bar chart

Description automatically generated

Figure 2. is a histogram of cell lengths in the 36659.jpg cell image. The 10 cells randomly chosen in Figure 1. were used and measured in um.

Takeaway

* The majority of the cells have cell lengths between 0 and 25 um. This indicates that shorter cells are more common.

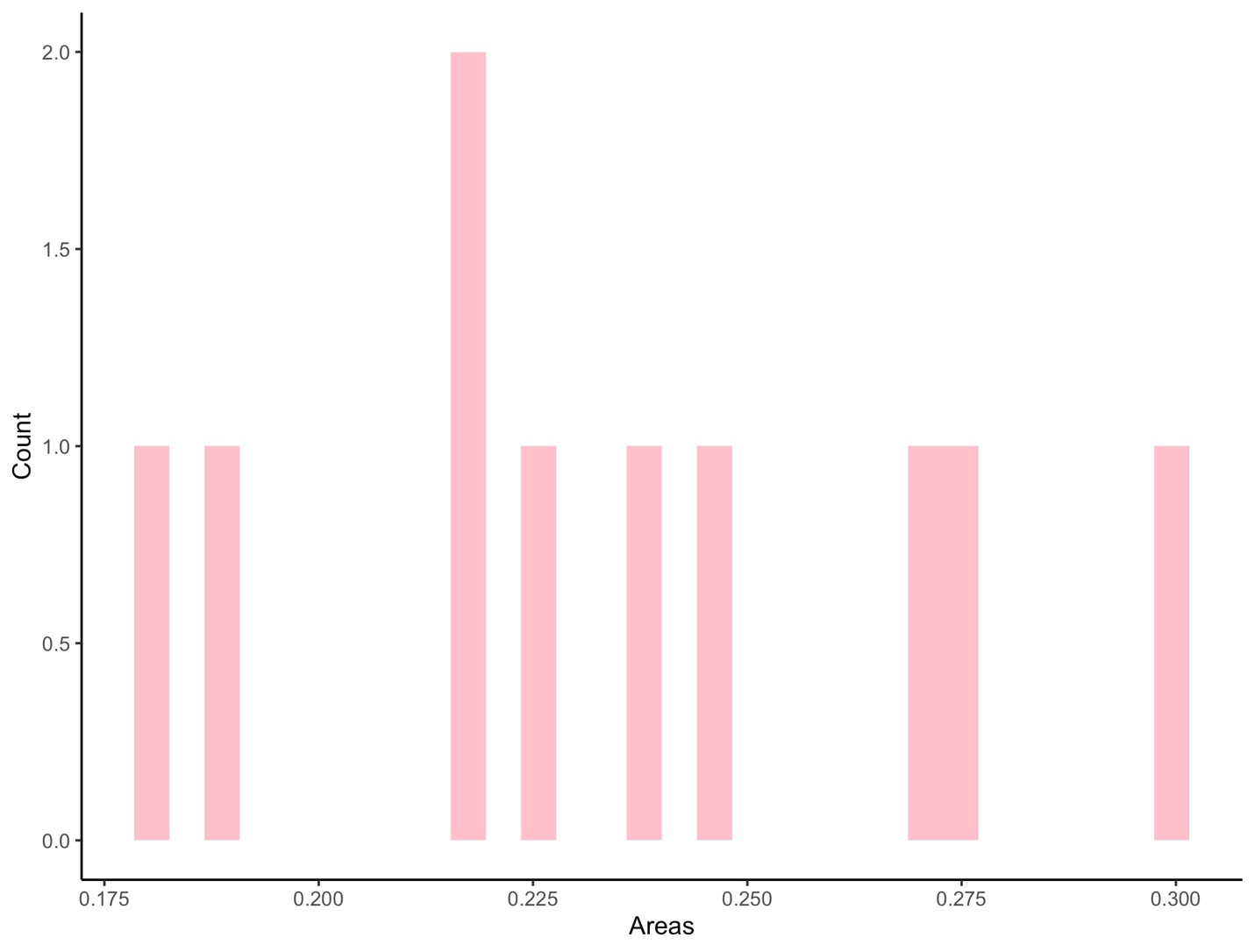


Figure 3. is a histogram of cell areas in the 36630.jpg cell image. 10 cells were randomly chosen and measured in um^2.

Takeaway

* The cells areas seem to be pretty spread out. This indicates that cells of all sizes are equally common.

Chart, bar chart

Description automatically generated

Figure 4. is a histogram of cell lengths in the 36630.jpg cell image. The 10 cells were randomly chosen in Figure 3. were used and measured in um.

Takeaway

* Most of the cells had lengths in the middle of the distribution. This indicates that cells between lengths 0.50 um to 0.65 um are more commonly found.

Chart

Description automatically generated

Figure 5. is a histogram of cell areas in the 19129.jpg cell image. 10 cells were randomly chosen and measured in um^2.

Takeaway

* The distribution of cell areas in this image seems to be evenly spread out. This implies that cells of low and high area are equally common.

Chart, histogram

Description automatically generated

Figure 6. is a histogram of cell lengths in the 19129.jpg cell image. The 10 cells randomly chosen in Figure 5. were used and measured in um.

Takeaway

* The majority of lengths fall in the middle and the end of the distribution. This indicates that cells of medium and large size are more common.