Lecture 4

William Cipolli

Colgate University

September 10, 2019

Homework for September 10-17

- Week 2
  - Diagnostic 1 was due Monday
  - Listen to the posted NPR podcast – due for today
- Week 3
  - Diagnostic 2 will be posted after class Thursdays 9/12 – due by noon 9/16
  - Complete Homework 0 (Posted) – due 11:59p 9/17

How To See The Future (No Crystal Ball Needed)

Q: What stuck out to you in this reading?

My Takeaways – How To See The Future (No Crystal Ball Needed)

- I had no idea about climate change (carbon from thawing permafrost)
- "Making someone angry isn't going to change their mind at all."
- "How do you get people to see there is value in understanding what you don't understand"
- Prediction has always been important both when heeded and ignored

My Takeaways – How To See The Future (No Crystal Ball Needed)

- We don't like uncertainties, and it doesn't suit us well to seem uncertain as predictors even if were are to some extent (Cassandra)
- Personal biases play into both how the predictions/warnings are formed and how they are perceived by others
  - We believe predictions that confirm our beliefs
  - We dismiss predictions that don't agree with our beliefs

My Takeaways – How To See The Future (No Crystal Ball Needed)

- Effective warnings are made easy to see, regardless of target beliefs, by showing "big" numbers and alarming visualization (this also gets us into trouble)
- Being an insider and having status makes the above easier
- Knowing how to read your prediction through the lens of others can help you make your prediction seem relevant to the target
- We can't ask people to go to far out of their comfort zone, so predictions need to be made easy to take action on or make a hard decision (government/NASA)

My Takeaways – How To See The Future (No Crystal Ball Needed)

- Be careful of predictions that make complicated things seem simple where there is no penalty for being wrong
- Being a good predictor often has much to do with pointing out the pitfalls of human thinking
- Pundits are more entertaining/interesting than correct – analysis, not real prediction
- "distinct possibility" and other "elastic phrases" (20-80%)
- Assigning probability is better, but more subject to embarrassment – "Hilary Clinton has a 75% chance to win" (Nate Silver)

My Takeaways – How To See The Future (No Crystal Ball Needed)

- Even when wrong, people look back and see things as they were right all along (hindsight bias)
- We don't check accuracy over time, which is necessary to weed out "lucky" predictors
- Having an accurate depiction of our prediction ability is important for improvement
- "Distinguishing more degrees of maybe" is an important skill
- We should update our forecasts frequently, making small adjustments

My Takeaways – How To See The Future (No Crystal Ball Needed)

- Zoom out – use historical data (the outside view)
- Wedding – how likely do you think it is that this couple is going to be married
  - Inside view – how we think:
    - at the wedding you might say 5% – it's a beautiful ceremony, the couple is so happy
  - Outside view – how predictors should think:
    - Socio-demographics of couple (parents/money/race/ethnicity)
    - Base rate of divorce
    - Characteristics of the couple
    - Have they been married before
    - etc

My Takeaways – How To See The Future (No Crystal Ball Needed)

- We are more easily persuaded by stories and metaphors – a flood in North America versus a flood in California caused by earthquake cracking a dam
- How do we distinguish between lucky and good? Streaky performance? Tournament of 16 – coin toss.
- You can never be 100% sure
- We're interested in big questions, not the narrow questions predictors answer (so often, the narrow answer is framed to answer the big question... often inappropriately)

My Takeaways – How To See The Future (No Crystal Ball Needed)

- We like specific number – 43% – which seem more certain
- People would be better off if they were more honest with themselves about their beliefs (does the evidence support it? are you cheering for your team?)

Wrapping up

Any questions about our basic programming skills in R?

- Objects
- Vectors
- Data frames
- Conditional Statements
- Loops
- Functions

Research Idea

- After exploring OKCupid data in an introductory statistics course, a student proposed their senior project in Psychology to explore race biases in dating online
- This is an area of racism that may have a harmful effect in a subtle way because people often justify racial prejudices as dating preferences

Research Questions

- How do ratings of interest in an online dating profile change once a participant learns the race of the person to whom the profile belongs?
- To which group of participants is race self-reported as least influential in their interest?
- Does learning that a high percentage of matches on a particular dating are mixed-race matches change any biases observed?

Data Collection

- Participants were recruited through Amazon Mechanical Turk
- Participants were asked questions about themselves and their demographic information
- Participants were asked to assess partial dating profiles (without race) and full dating profiles (with race)

Overview

Overview and Procedure:

You are being asked to take part in a research study. To join the study is voluntary. You may refuse to join, or you may withdraw your consent to be in the study, for any reason, at any time, without penalty. Details about this study are discussed below. It is important for you to understand this information so that you can make an informed choice about being in this research study.

This study examines evaluations of online dating profiles. For that reason, only people who currently identify as "dating" or "single" will be eligible for this study. Your participation in this study will last approximately 10-15 minutes.

Risks

Risks and Benefits:

We anticipate few risks in this study. Beyond compensation, you may not benefit personally from being in this study.

Confidentiality

Confidentiality:

By completing our study through Amazon Mechanical Turk (MTurk) you are subject to our own standards of confidentiality (the researchers posting this study) as well as those of Amazon.com.

Our standards of confidentiality: The only identifying information that we will collect will be your Mturk worker ID which is linked to your Amazon.com public profile. This ID will only be used to help us assign credit for study completion. Once data collection is complete and you have been paid for your participation we will delete all Mturk IDs from our data files. Thus, only completely de-identified data files will be shared among researchers once data collection is complete and you have been paid for your participation.

Amazon.com standards of confidentiality: Amazon.com has stated that the MTurk platform is NOT meant to support participant anonymity. MTurk worker IDs are linked to Amazon.com public profiles. Amazon.com may disclose information about you as an Mturk worker. Additionally, worker information may be available to others (who submit a request) for tax reporting purposes.

Payment and Rights

You will receive $.70 for participating in this study. There are no costs associated with being in the study.

Your Rights:

To join the study is voluntary. You may refuse to join. If you choose to join, you may also withdraw from this study at any time and without penalty.

First Questions

- ○ I am at least 18 years old and agree to participate in this study.

What is your relationship status?

- ○ Single
- ○ Dating
- ○ In a committed, monogamous relationship

Are you currently interested in dating?

- ○ Yes
- ○ Maybe
- ○ No

Eligible Participants

- We're looking for a sample that is representative of online daters
- Anyone who answers that they're in a monogamous relationship isn't eligible for the study
- Anyone who answers that they're not interested in dating isn't eligible for the study

Ethical Misleading

We tell participants about the purpose of this study – kind of – to motivate them to answer the questions about their dating preferences.

This study is being conducted by a team interested in improving the experience of people on a particular dating website (that will remain unnamed for privacy reasons). This dating website is trying to create an algorithm to best match people into successful relationships.

General Info

Dating websites and applications vary in how much information is included in people's profiles. Some applications (e.g., Tinder, Bumble) provide very little profile information.

Other websites (e.g., Match.com, eharmoney) include profiles that are much more detailed.

As a part of this research, we are interested in how much information is helpful to users when evaluating potential partners.

General Info

On the next several pages you will see real profiles of users of our website (photos omitted). First we will provide you with a version of their profile that is limited in information.

We will then ask you to report your reactions to the profile as honestly as possible, as if you were using this website and considering a date with the person.

General Info

Next we will show you the user's FULL profile (which includes a bit more information, but still omitting the photo to protect their identities).

We will again ask you to report your reactions to that profile.

General Info

Our hope with this research is to improve our dating website to improve user satisfaction and help people meet long-term matches.

We are particularly interested in how much information is desirable when reviewing potential mates before meeting in person.

General Info

First, please report your dating preference below so we can choose profiles from our database that match your preferences. While we know that some people may not care about the gender of a potential partner, for the purposes of this study, please indicate the type of profiles that you would like to examine.

- ○ Interested in viewing women's profiles (i.e., interest in dating women)
- ○ Interested in viewing men's profiles (i.e., interest in dating men)

Deciding Flow of the Survey

The question on the previous slide allows us to send the participant to the appropriate survey of ten different flows created from 5 example profiles.

- Survey 1: Asian (1), Black (5), White (4) – Female
- Survey 2: Asian (2), Black (1), White (5) – Female
- Survey 3: Asian (3), Black (2), White (1) – Female
- Survey 4: Asian (4), Black (3), White (2) – Female
- Survey 5: Asian (5), Black (4), White (3) – Female

Deciding Flow of the Survey

The question on the previous slide allows us to send the participant to the appropriate survey of ten different flows created from 5 example profiles.

- Survey 6: Asian (1), Black (5), White (4) – Male
- Survey 7: Asian (2), Black (1), White (5) – Male
- Survey 8: Asian (3), Black (2), White (1) – Male
- Survey 9: Asian (4), Black (3), White (2) – Male
- Survey 10: Asian (5), Black (4), White (3) – Male

Profile 1

About Me:

I was born and raised in Texas, so I absolutely love warm weather. I love going swimming, hanging by the beach, and just about anything that involves water. I have a job that I really enjoy, but I also savor time when I can relax. I'm a pretty good chef and I love cooking and baking for friends and family. My biggest fear is that my oven will break. I'm only half kidding.

Gender:
Woman
Race:
Asian
Hobbies:
Swimming, watching live sports
Favorite food:
Grilled cheese and tomato soup
My friends would say I'm:
A good listener and a very proud new aunt

Profile 2

About Me:

I would describe myself as fun-loving and a big fan of dogs. I have a young black lab who I love going on runs with. Sometimes I bring my dog to work with me, which luckily I am able to do in my job! (Side note: I think it scores me points with my boss, who is also crazy about dogs.) I'm also a huge fan of watching Netflix shows, trying new foods, and spending time with friends and family.

Gender:
Woman
Race:
Asian
Hobbies:
Running, reading mystery books, and watching movies
Favorite food:
Gourmet pizza
My friends would say I'm:
An easy-going person who laughs harder at my own jokes than other people do...

Profile 3

About Me:

My ideal weekend involves going to do something artsy (a play, a movie, a museum), eating good food, or doing something active outside. (Or, all of the above?) I also love to travel, largely because, when I was young, I moved around a lot because my dad was a pilot. That said, I am also a master of doing absolutely nothing and enjoying the moment...or at least I try to be!

Gender:
Woman
Race:
Asian
Hobbies:
Cooking, sports
Favorite food:
Any type of cheese (really, any type...)
My friends would say I'm:
A great conversationalist

Profile 4

About Me:

I really like to sing (warning: not well!). But, that doesn't keep me from belting out songs in the shower. Whether it's watching a movie, grabbing drinks, or relaxing in the pool, I'm often spending time with friends and family. During the week, I spend a fair amount of time working, but also enjoy relaxing at night with people I care about. Luckily, I really enjoy my job because I get to interact with people from different backgrounds on a daily basis.

Gender:
Woman
Race:
Asian
Hobbies:
Painting or anything artistic, singing (not well), dancing (not well)
Favorite food:
Anything Italian!
My friends would say I'm:
A good listener who also (tries to) give good advice

Profile 5

About Me:

You know that person in your group of friends who is always planning something? Well, that's me. I love planning fun events either for my friends, family, or myself! My perfect weekend would include a hike, followed by a visit to a new local brewery or restaurant. I'm also an avid pie baker and swimmer, but have yet to master doing both at the same time. But you can watch me try! Just kidding, that actually sounds like a bad idea.

Gender:
Woman
Race:
Asian
Hobbies:
Baking, swimming, and watching comedies
Favorite food:
Anything I haven't tried before—I love new things
My friends would say I'm:
Down-to-earth and sarcastic

Example Flow – Asian Woman

On the next page, you will see an **abbreviated user profile.** Please read the profile carefully, imagine you are using the site to find a potential match, and report your evaluations of the profile.

Example Flow – Asian Woman

About Me:

I was born and raised in Texas, so I absolutely love warm weather. I love going swimming, hanging by the beach, and just about anything that involves water. I have a job that I really enjoy, but I also savor time when I can relax. I'm a pretty good chef and I love cooking and baking for friends and family. My biggest fear is that my oven will break. I'm only half kidding.

Example Flow – Asian Woman

Please answer the following in response to this **partial profile.** When answering the following questions, please remember to imagine you are actively using our dating website in order to meet a romantic partner, and that you come across this profile.

Strongly Disagree Strongly Agree

0 10 20 30 40 50 60 70 80 90 100

Example Flow – Asian Woman

Participants will use a slider to input their agreement on the following questions:

- I find this person appealing based on the information in their profile.
- I do not find this person appealing based on the information in their profile.
- I am romantically interested in this person based on the information in their profile.

Example Flow – Asian Woman

How likely would you be to consider **going on a date** with the person whose partial profile you just read?

Very Unlikely Very Likely

0 10 20 30 40 50 60 70 80 90 100

Interest

We calculate interest as

$$\text{interest} = \frac{\text{appeal} + (1 - \text{not appeal}) + \text{romantic} + \text{date}}{4}$$

Example Flow – Asian Woman

On the next page, you will see the **full user profile** (rather than the abbreviated version). Please read the profile carefully, imagine you are using the site to find a potential match, and report your evaluations of the profile.

Example Flow – Asian Woman

About Me:

I was born and raised in Texas, so I absolutely love warm weather. I love going swimming, hanging by the beach, and just about anything that involves water. I have a job that I really enjoy, but I also savor time when I can relax. I'm a pretty good chef and I love cooking and baking for friends and family. My biggest fear is that my oven will break. I'm only half kidding.

Gender:
Woman
Race:
Asian
Hobbies:
Swimming, watching live sports
Favorite food:
Grilled cheese and tomato soup
My friends would say I'm:
A good listener and a very proud new aunt

Example Flow – Asian Woman

Please answer the following in response to this **full profile**. When answering the following questions, please remember to imagine you are actively using our dating website in order to meet a romantic partner, and that you come across this profile.

Strongly Disagree Strongly Agree

0 10 20 30 40 50 60 70 80 90 100

Example Flow – Asian Woman

Participants will use a slider to input their agreement on the following questions:

- I find this person appealing based on the information in their profile.
- I do not find this person appealing based on the information in their profile.
- I am romantically interested in this person based on the information in their profile.

Example Flow – Asian Woman

How likely would you be to consider **going on a date** with the person whose full profile you just read?

Very Unlikely Very Likely

0 10 20 30 40 50 60 70 80 90 100

Example Flow – Asian Woman

Participants were then asked to rate the information on the profile in whether you personally felt it was a "Turn On," "Turn Off," or "Neutral/Neither"

- I was born and raised in Texas, so I absolutely love warm weather.
- I love swimming, surfing, and just about anything that involves water.
- I have a job that I really enjoy, but I also savor time when I can relax.
- I'm a pretty good chef and I love cooking and baking.
- My biggest fear is that my oven will break. I'm only half kidding.

Example Flow – Asian Woman

Participants were then asked to rate the information on the profile in whether you personally felt it was a "Turn On," "Turn Off," or "Neutral/Neither"

- Hobbies: Swimming, watching live sports
- Favorite food: Grilled cheese and tomato soup
- Gender: Woman
- Race: Asian
- My friends would say I'm: A good listener and a very proud new aunt

Example Flow – Asian Woman

Participants were then asked to rank the following components of the profile based on the degree to which they increased how attracted you were to the profile from 1 (most influential at increasing my attraction) to 6 (least influential at increasing my attraction).

- About me
- Hobbies
- Favorite Food
- Race
- Gender
- What their friends would say about them

Example Flow – ...

Participants then completed the same tasks for a Black profile and a White profile.

Example Flow – General Dating Beliefs

How important is it for you to be **the same as your romantic partner** in the following characteristics?

Not important at all — Very Important

0 10 20 30 40 50 60 70 80 90 100

Example Flow – General Dating Beliefs

- Importance of having the same hobbies and interests as my partner
- Importance of being the same race as my partner
- Importance of being the same religious affiliation as my partner
- Importance of being the same political orientation as my partner

Example Flow – General Dating Belief

Please rate how much you agree with the following statements.

Strongly Disagree — Strongly Agree

0 10 20 30 40 50 60 70 80 90 100

Example Flow – General Dating Beliefs

- Most romantic matches are between people with the same religious affiliation.
- Romantic matches are likely to last longer when partners have the same religious affiliation.
- Most romantic matches are between people with the same political affiliation.
- Romantic matches are likely to last longer when partners have the same political affiliation.

Example Flow – General Dating Beliefs

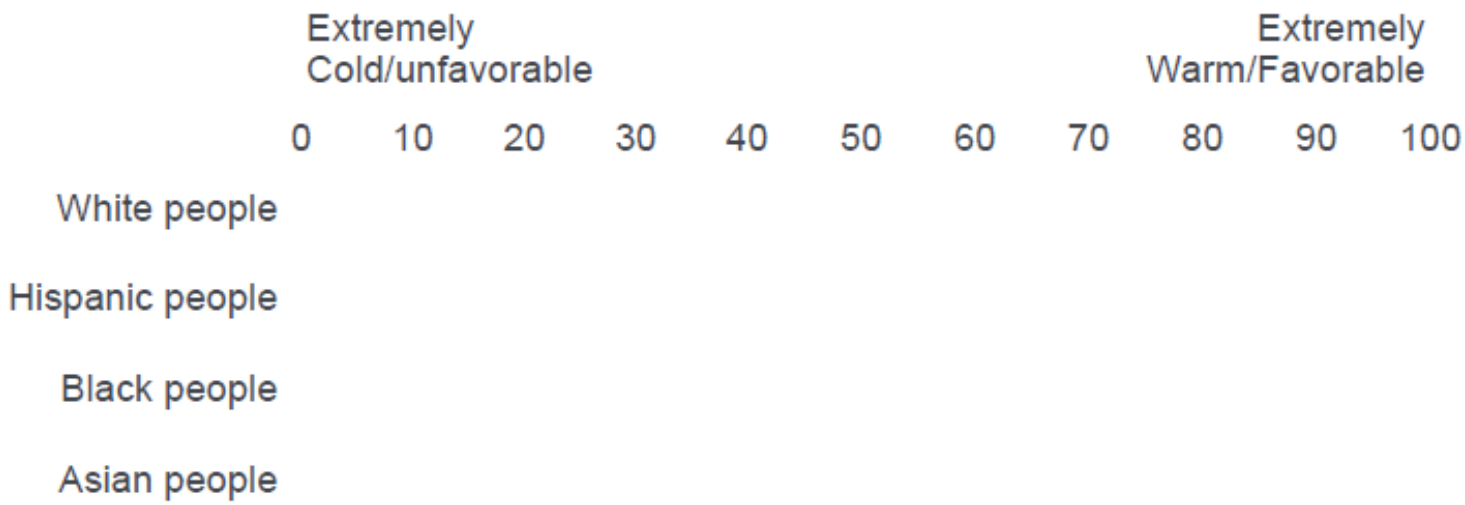
- Most romantic matches are between people of the same race.
- Romantic matches are likely to last longer when the partners are the same race.
- Most romantic matches are between people with the similar hobbies and interests.
- Romantic matches are likely to last longer when partners have the same hobbies and interests.

Example Flow – Feeling Thermometer

Now, we'd like to get your feelings toward certain groups. Below are names of groups. Using the scale below, we'd like you to rate how warm or cold you feel toward each group. Ratings between 50 degrees and 100 degrees mean that you feel favorable and warm toward the group. Ratings between 0 degrees and 50 degrees mean that you don't feel favorable toward the group and that you don't care too much for that group. You would rate the group at the 50 degree mark if you don't feel particularly warm or cold toward the group.

Example Flow – Feeling Thermometer

How favorable or unfavorable do you feel toward...



Example Flow – Egalitarian Motivation

Please indicate your agreement to the following items.

	Strongly Disagree 0	10	20	30	40	50	60	70	80	90	Strongly Agree 100
In my own life, I try to behave in ways that promote racial equality.											
I go out of my way to try to increase harmony between races.											

Example Flow – Race

The following questions concern various reasons or motivations people might have for trying to respond in nonprejudiced ways toward **Black people**. Some of the reasons reflect internal--personal motivations, whereas others reflect more external--social motivations. Of course, people may be motivated for both internal and external reasons; we want to emphasize that neither type of motivation is by definition better than the other. In addition, we want to be clear that we are not evaluating you or your individual responses. All of your responses will be completely confidential. We are simply trying to get an idea of the types of motivations that people in general have for responding in nonprejudiced ways. It is important that you respond to each of the questions openly and honestly. Please click on the response that best represents your motivations. There are 5 items total.

Example Flow – Race

Please indicate your agreement to the following items.

Strongly Disagree — Strongly Agree

0 10 20 30 40 50 60 70 80 90 100

Example Flow – Race

- I attempt to act in non-prejudiced ways toward Black people because it is personally important to me.
- According to my personal values, using stereotypes about Black people is okay.
- I am personally motivated by my beliefs to be non-prejudiced toward Black people.
- Because of my personal values, I believe that using stereotypes about Black people is wrong.
- Being non-prejudiced toward Black people is important to my self-concept.

Example Flow – Race

- Because of today's PC (politically correct) standards I try to appear no prejudiced toward Black people.
- I try to hide any negative thoughts about Black people in order to avoid negative reactions from others.
- If I acted prejudiced toward Black people, I would be concerned that others would be angry with me.
- I attempt to appear nonprejudiced toward Black people in order to avoid disapproval from others.
- I try to act nonpredjudiced toward Black people because of pressure from others.

Example Flow – Race

The following questions concern various reasons or motivations people might have for trying to respond in nonprejudiced ways toward **Asian people**. Some of the reasons reflect internal--personal motivations, whereas others reflect more external--social motivations. Of course, people may be motivated for both internal and external reasons; we want to emphasize that neither type of motivation is by definition better than the other. In addition, we want to be clear that we are not evaluating you or your individual responses. All of your responses will be completely confidential. We are simply trying to get an idea of the types of motivations that people in general have for responding in nonprejudiced ways. It is important that you respond to each of the questions openly and honestly. Please click on the response that best represents your motivations. There are 5 items total.

Example Flow – Race

Participants then answered the same 10 questions about Asian people.

Example Flow – Demographics

The study is almost complete. On the following pages we will ask you several final questions.

How old are you?

What is your gender?

- Man
- Woman
- Other

Example Flow – Demographics

What is your ethnicity?

- ○ Hispanic or Latino
- ○ Not Hispanic or Latino

Example Flow – Demographics

What is your race? (check all that apply)

- [ ] White or Caucasian
- [ ] Black or African American
- [ ] Native American or Pacific Islander
- [ ] Asian
- [ ] Other

When you selected "other" for your race, what did you mean?

Example Flow – Demographics

Which best describes your sexual orientation?

- ○ Gay/Lesbian
- ○ Heterosexual
- ○ Bisexual
- ○ Asexual
- ○ Pansexual
- ○ Other

How would you define your sexual orientation?

Example Flow – Demographics

Think of this ladder as representing where people stand in the United States.

At the top of the ladder (rung 10) are the people who are the best off – those who have the most money, the most education and the most respected jobs.

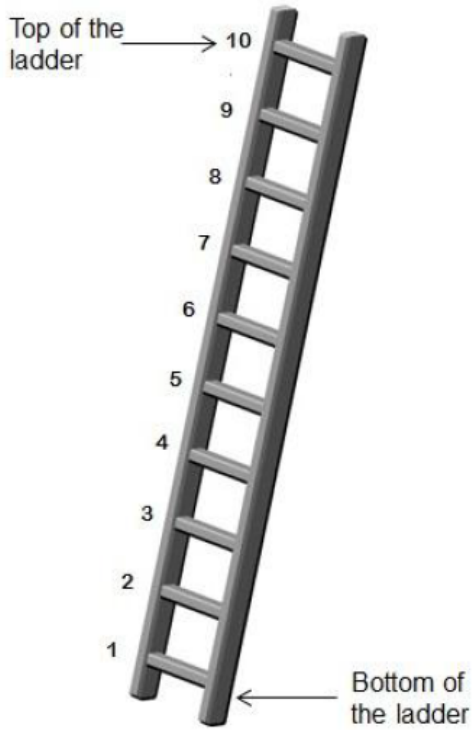
At the bottom of the ladder (rung 1) are the people who are the worst off – those who have the least money, least education, and the least respected jobs or no job.

The higher up you are on this ladder, the closer you are to the people at the very top; the lower you are, the closer you are to people at the very bottom.

Example Flow – Demographics

Where would you place yourself on this ladder?

Click the number of the rung where you think you stand at this time in your life, relative to other people in the United States.



1	2	3	4	5	6	7	8	9	10
Bottom of the ladder	○	○	○	○	○	○	○	○	Top of the ladder

Example Flow – Demographics

Participants were then asked to select the category they belong to each of the following questions.

- What is your yearly household income?
- What is your highest level of education?
- Did you take this study on a smartphone, computer, or tablet?

Example Flow – Debrief

What do you think the purpose of the current study is?

Did you see any connections between the data about matches made on our website and the profile ratings we asked you to make later in the study?

- ○ Yes
- ○ No

Example Flow – Debrief

What did you think was the connection between the data about matches made on our website and the profile ratings we asked you to make later in the study? Please describe in detail, as this will help your data be more meaningful. Thank you.

Example Flow – Debrief

On the next page you will view more information about the purpose of the study.

Your completion code is 3399.

Example Flow – Debrief

Data from dating websites indicates that White people tend to prefer dating other White people (Mendelsohn, Shaw, & Chesire, 2014). In the present study, we try to better understand these racebased preferences and also test a potential intervention to reduce these racially-biased dating selections.

Example Flow – Debrief

If you were in the first sample, you made ratings of dating profiles. In particular, you viewed dating profiles with and without race information included. Some of these dating profiles were described as written by White people, some by Black people, and some by Asian people. These were not real profiles, but instead were created to seem like real profiles. We were interested in how preferences for the profiles would change with and without race information.

Example Flow – Debrief

If you were in the second sample, you completed the same ratings of the dating profiles. However, before you made these ratings you also read an article about race being biologically determined and fixed, an article about race being fluid and changeable, or a scientific article unrelated to race. These were not real scientific articles. Instead they were intended to shift the way you might think about race. We hypothesized that learning that race is flexible might lead people to be more open to dating people of a different race than their own. For a recent, real scientific article that discusses the genetic basis of race see the following resource:

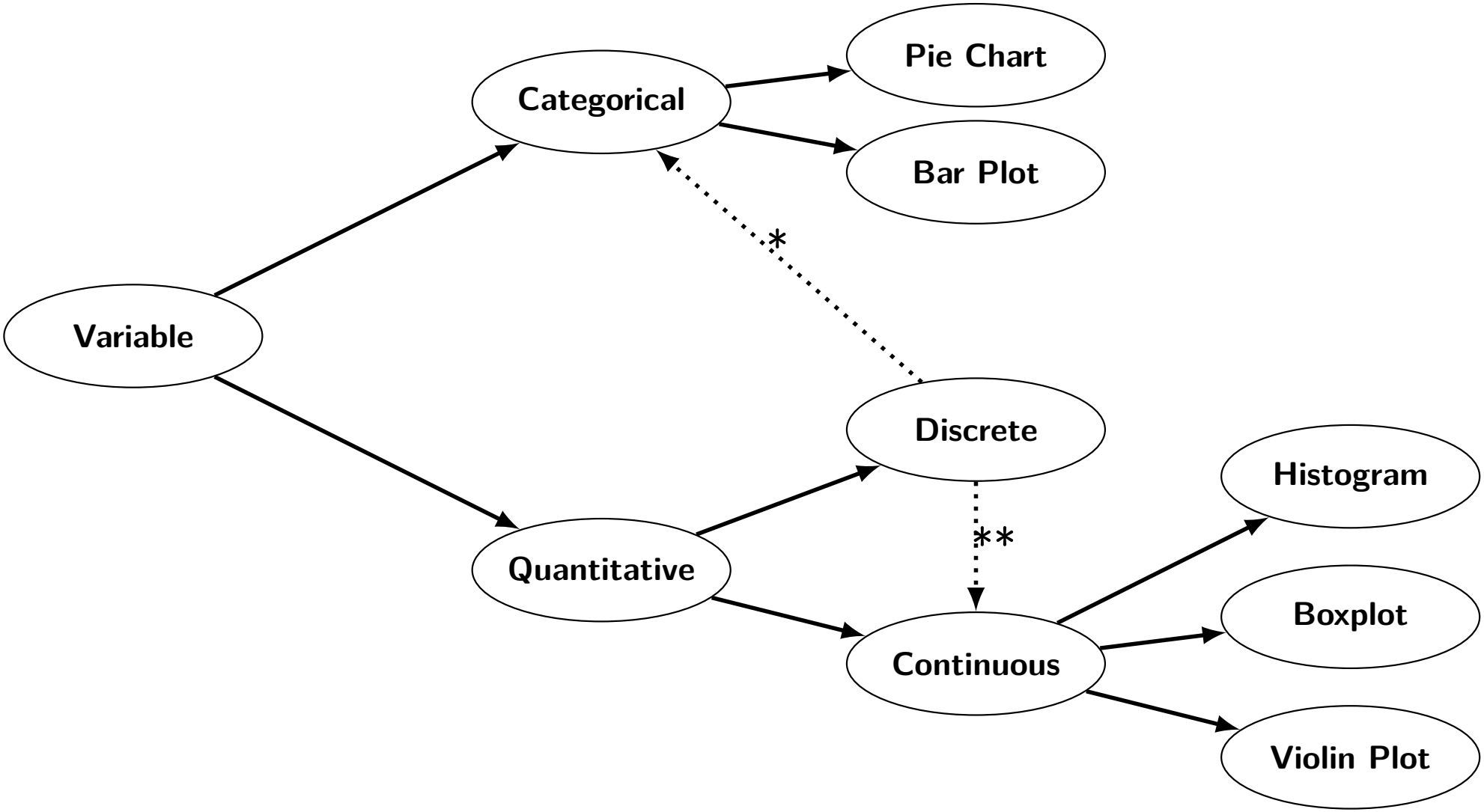
Example Flow – Debrief

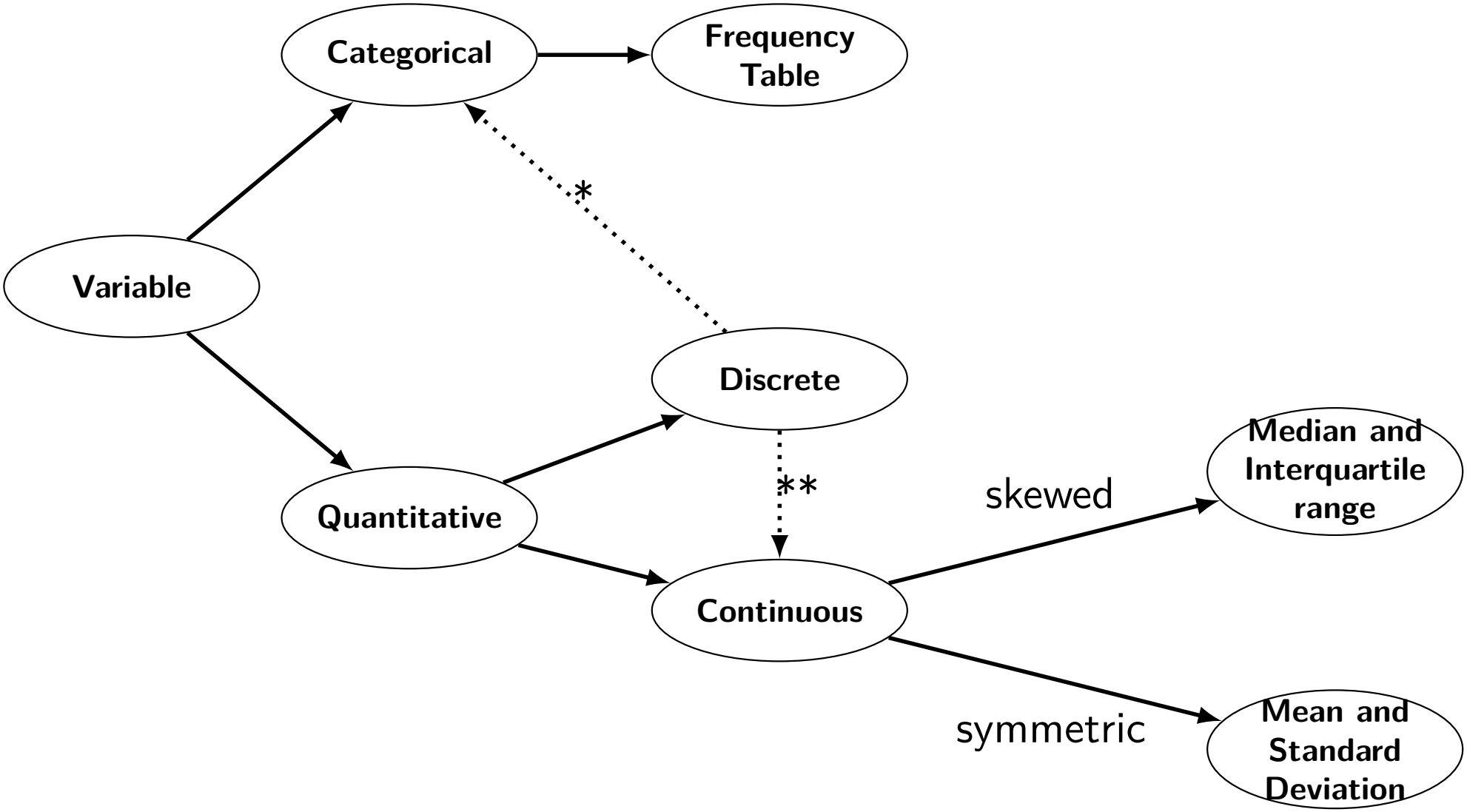
Yudell, M., Roberts, D., DeSalle, R., & Tishkoff, S. (2016). Taking race out of human genetics. Science, 351(6273), 564-565.
Our hope with this research is to better understand how race, and racial attitudes, influence the expression of dating preferences.
If you would like to talk with a mental health professional, http://blahtherapy.com/ provides low-cost option for online therapy.
Should you have questions regarding your rights as a participant in research, please contact the Institutional Review Board at IRB_Chair@psych.colgate.edu. Should you have questions about the study, please feel free to contact the researchers working on this project:

Random Variables

A **variable** is a characteristic or measurement that we have for each observation.

- **Categorical** $\longrightarrow$ places observations into one of several groups or categories
  - **Nominal** $\longrightarrow$ assumes groups are made up of named categories
  - **Ordinal** $\longrightarrow$ assumes groups are made up of ordered categories
- **Quantitative** $\longrightarrow$ assumes numerical values
  - **Continuous** $\longrightarrow$ assumes the variable is double measured
  - **Discrete** $\longrightarrow$ assumes the variable is countable


Variable
Categorical
Quantitative
Pie Chart
Bar Plot
*
Discrete
**
Continuous
Histogram
Boxplot
Violin Plot


Categorical
Frequency Table
*
Variable
Discrete
**
Quantitative
skewed
Median and Interquartile range
Continuous
symmetric
Mean and Standard Deviation

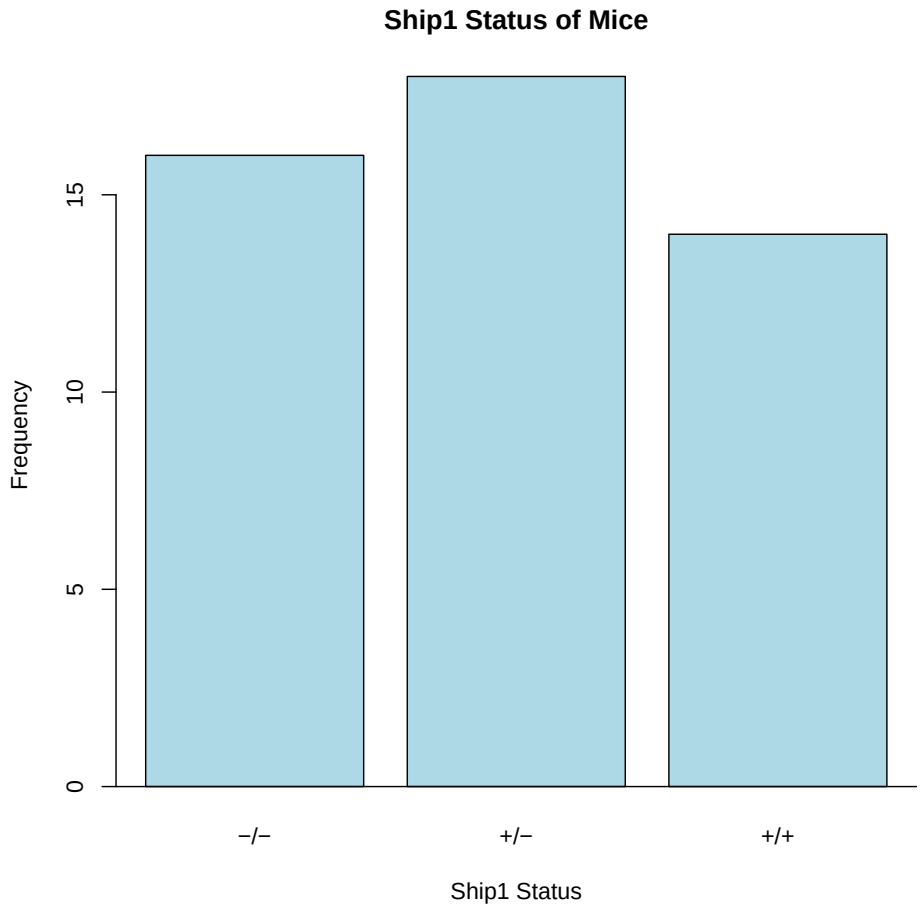
Numerical Summary

Operator	Functionality
table(...)	creates a frequency table
prop.table(...)	creates a relative frequency table

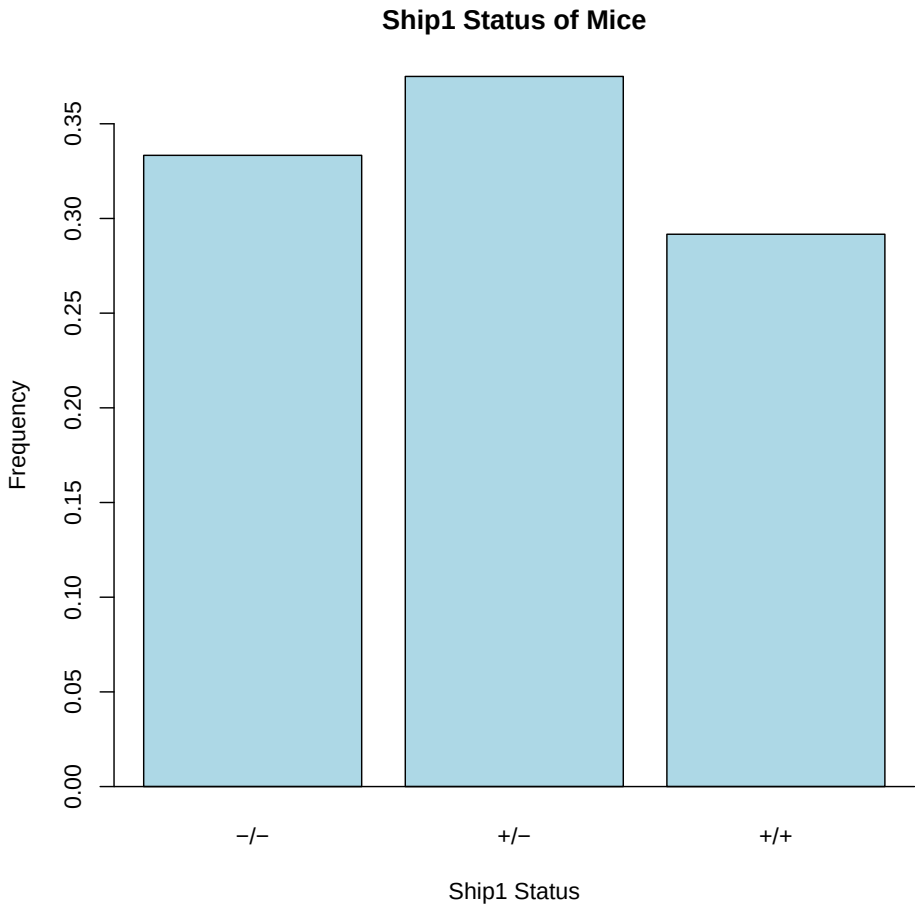
Graphical Summary – Barplot

Operator	Functionality
par(mfrow=c(r,c))	makes the graphics environment a $r \times c$ grid
barplot(...)	creates a barplot for a table of data
abline(h=0)	draws a horizontal line at $y = 0$

Graphical Summary – Barplot



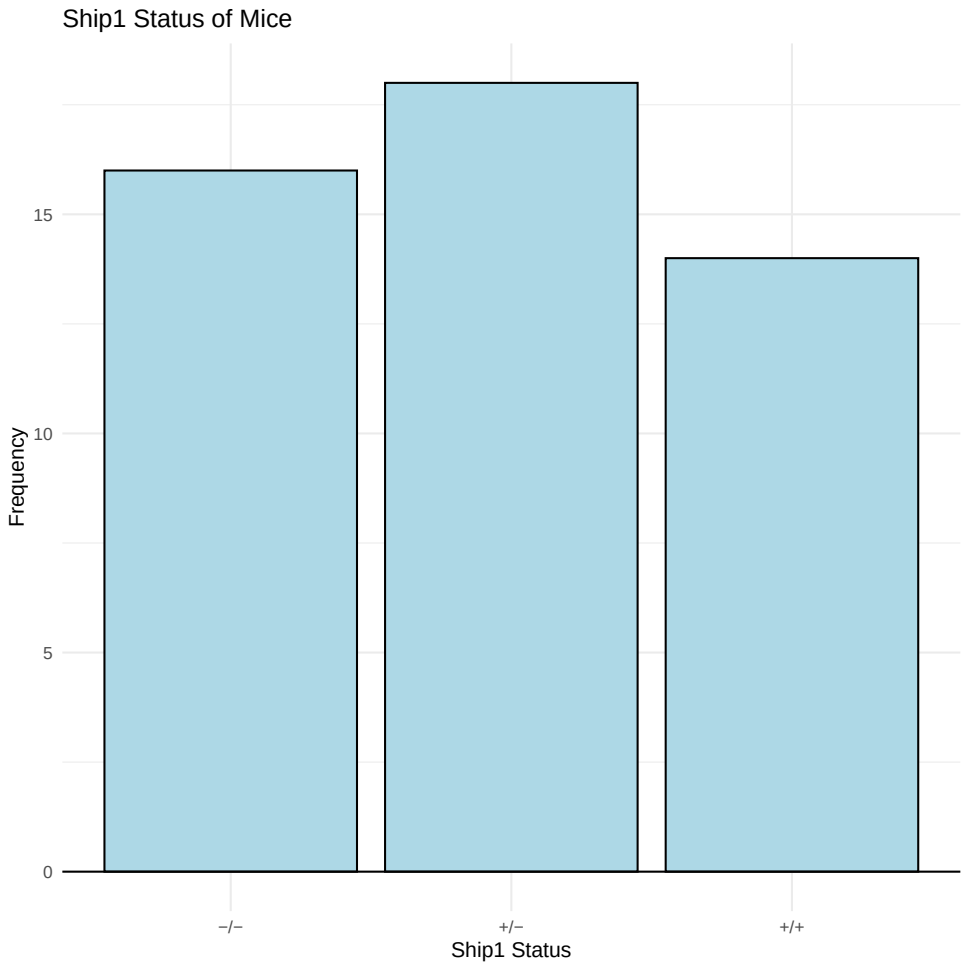
Graphical Summary – Barplot



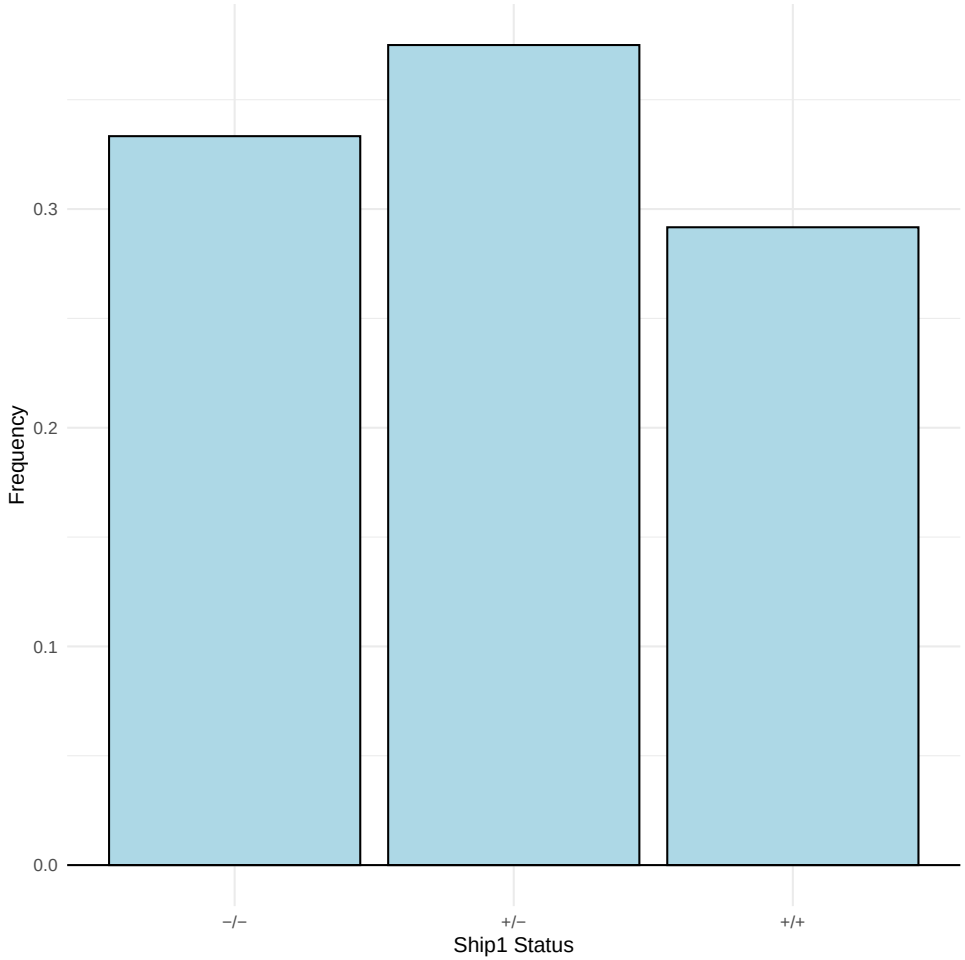
Graphical Summary – Barplot (ggplot)

Operator	Functionality
ggplot(...)	denotes a ggplot command has started
aes(...)	describe how variables in the data are mapp
geom_bar(...)	creates a barplot in ggplot
xlab(...)	edit x axis label in ggplot
ylab(...)	edit y axis label in ggplot
ggtitle(...)	edit title in ggplot
geom_hline(yintercept=0)	draws a horizontal line at $y = 0$
theme_minimal(...)	removes grey background of ggplot
grid.arrange(...)	arranges multiple ggplots onto one screen

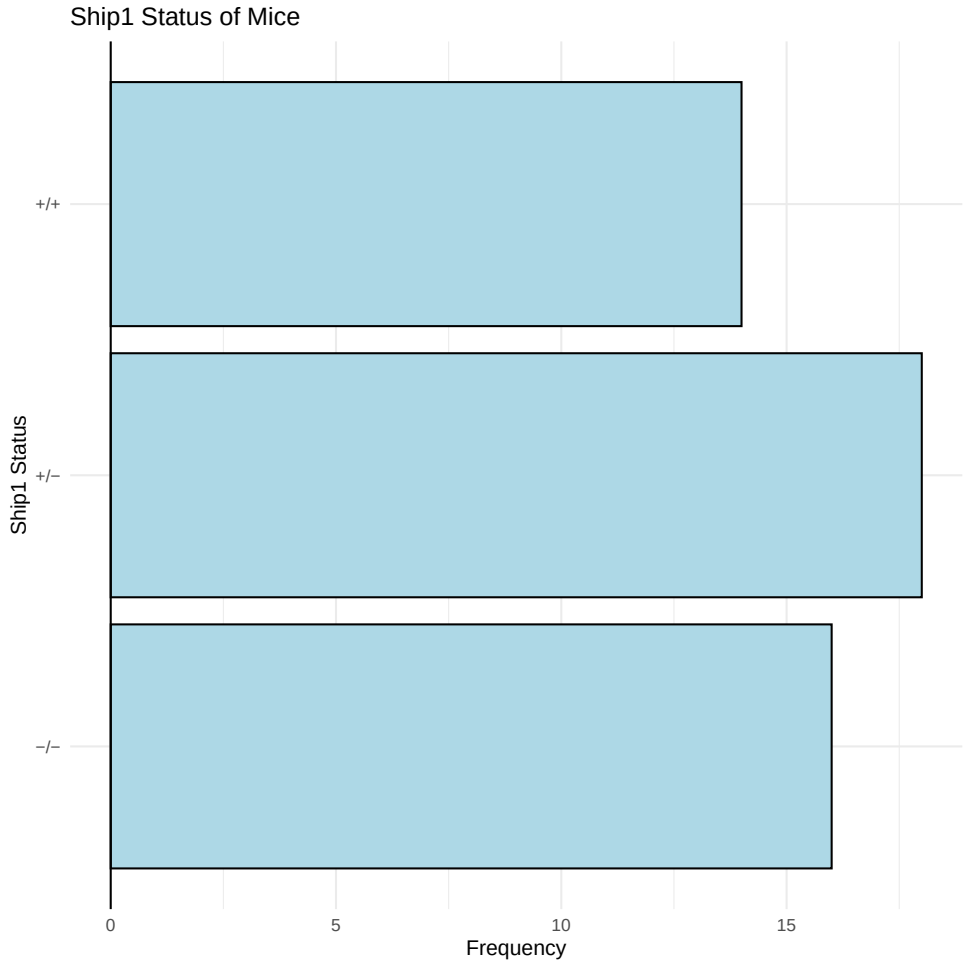
Graphical Summary – Barplot



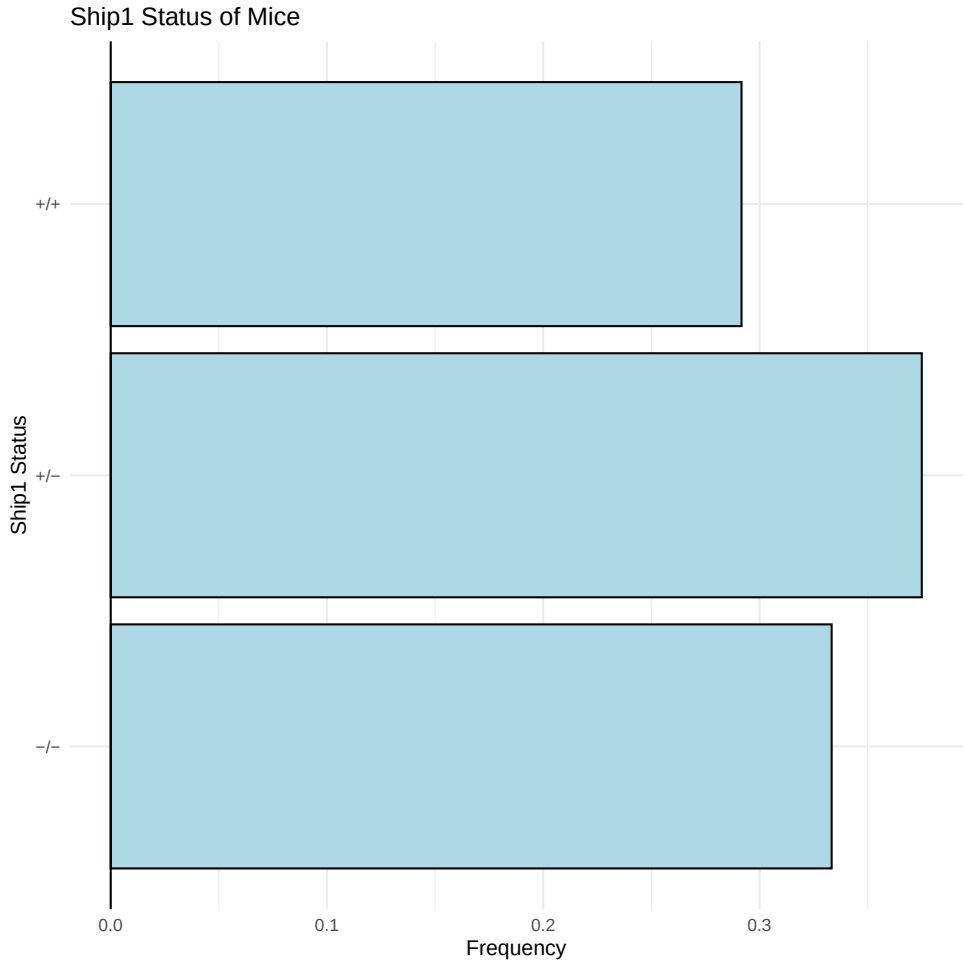
Graphical Summary – Barplot



Graphical Summary – Barplot



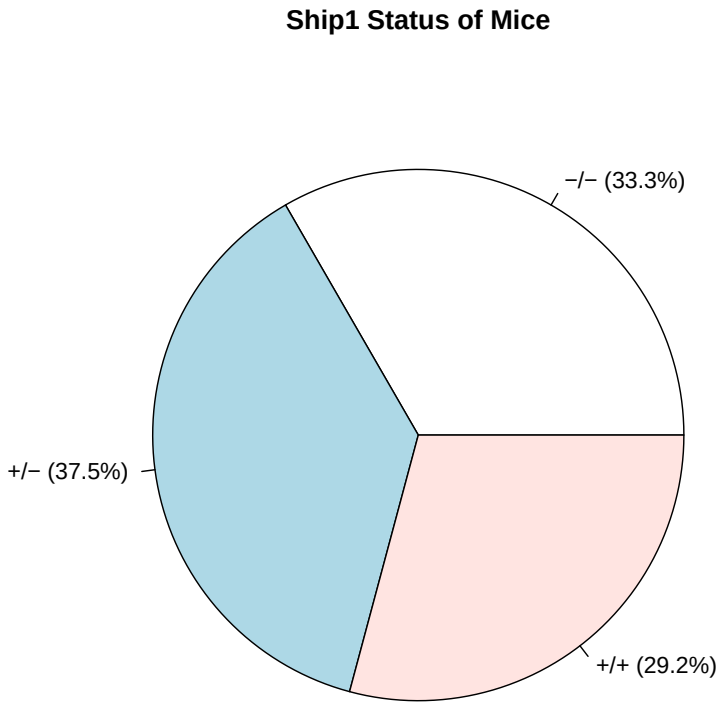
Graphical Summary – Barplot



Graphical Summary – Pie Chart

Operator	Functionality
pie(...)	creates a pie chart for a table of data

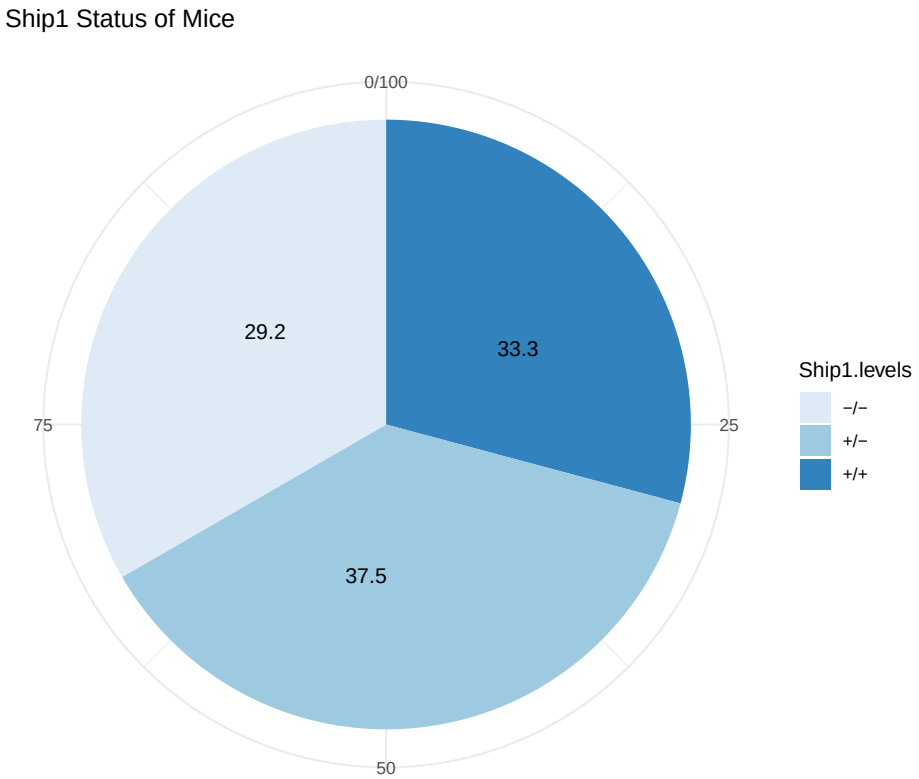
Graphical Summary – Pie Chart



Graphical Summary – Pie Chart (ggplot)

Operator	Functionality
coord_polar(...)	creates a pie chart for a bar plot in ggplot
scale_fill_brewer(...)	specify color palate for ggplot

Graphical Summary – Pie Chart



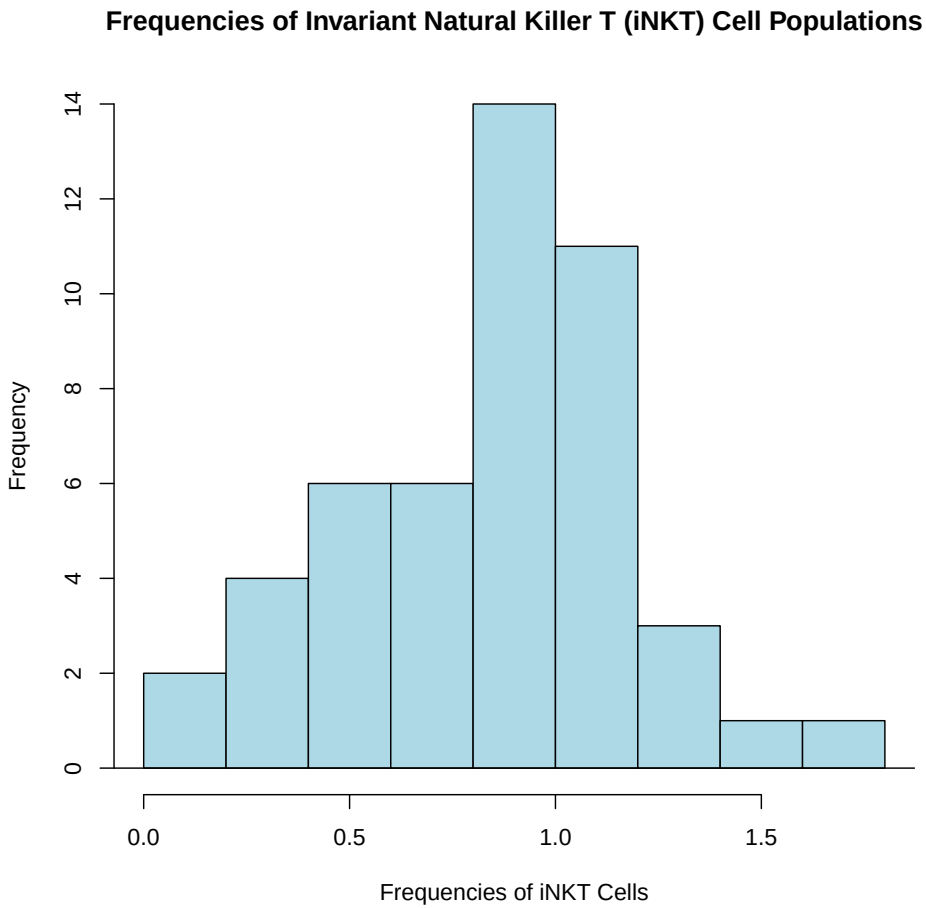
Numerical Summary

Operator	Functionality
mean(...)	provides the mean for a vector
sd(...)	provides the standard deviation for a vector
min(...)	provides the minimum for a vector
quantile(...)	provides the specified percentile for a vector
median(...)	provides the median for a vector
max(...)	provides the max for a vector
summary(...)	provides the five number summary and mean for a vecto
IQR(...)	provides the interquartile rang for a vector

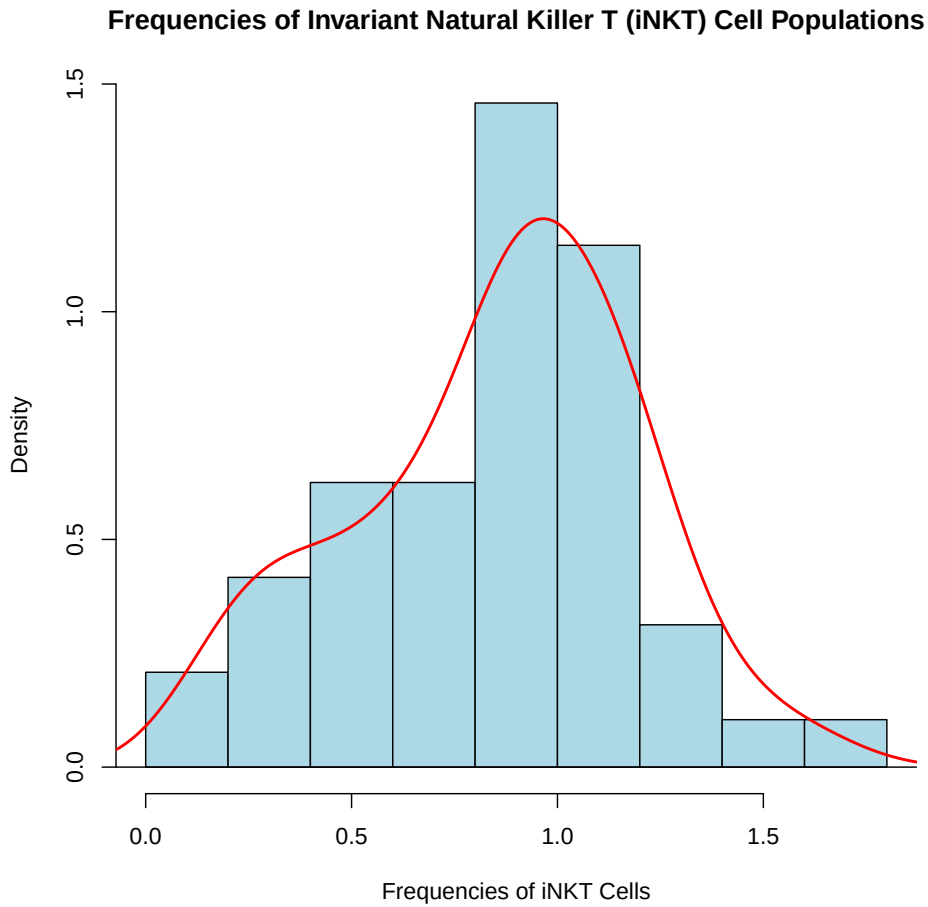
Graphical Summary – Histogram

Operator	Functionality
hist(...)	creates a histogram for a vector of data
lines(...)	adds a line to a plot
density(...)	computes kernel density estimate for a vector of data
legend(...)	creates a legend for a plot

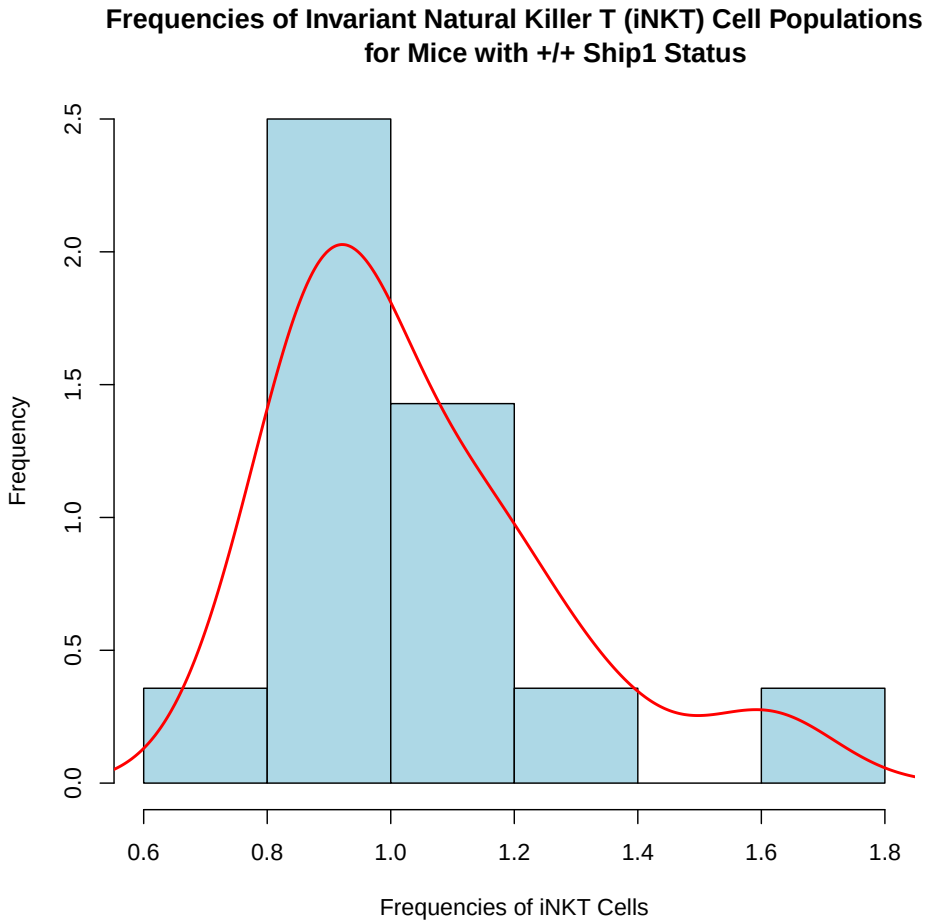
Graphical Summary – Histogram



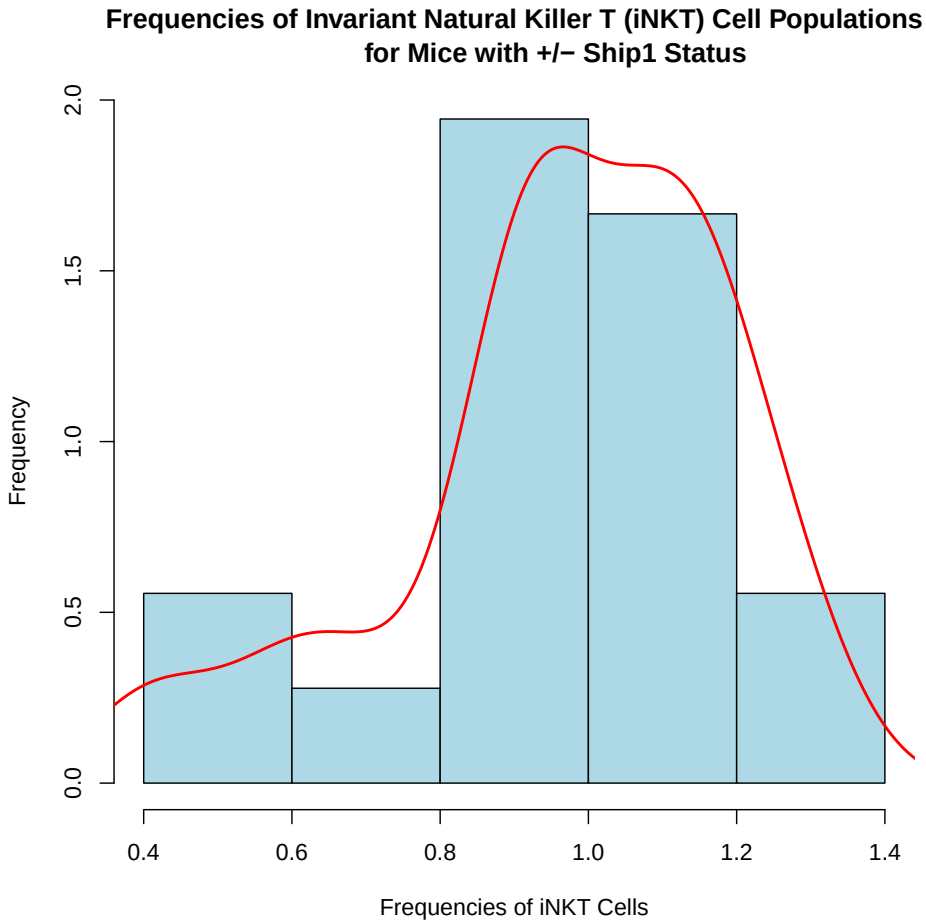
Graphical Summary – Histogram



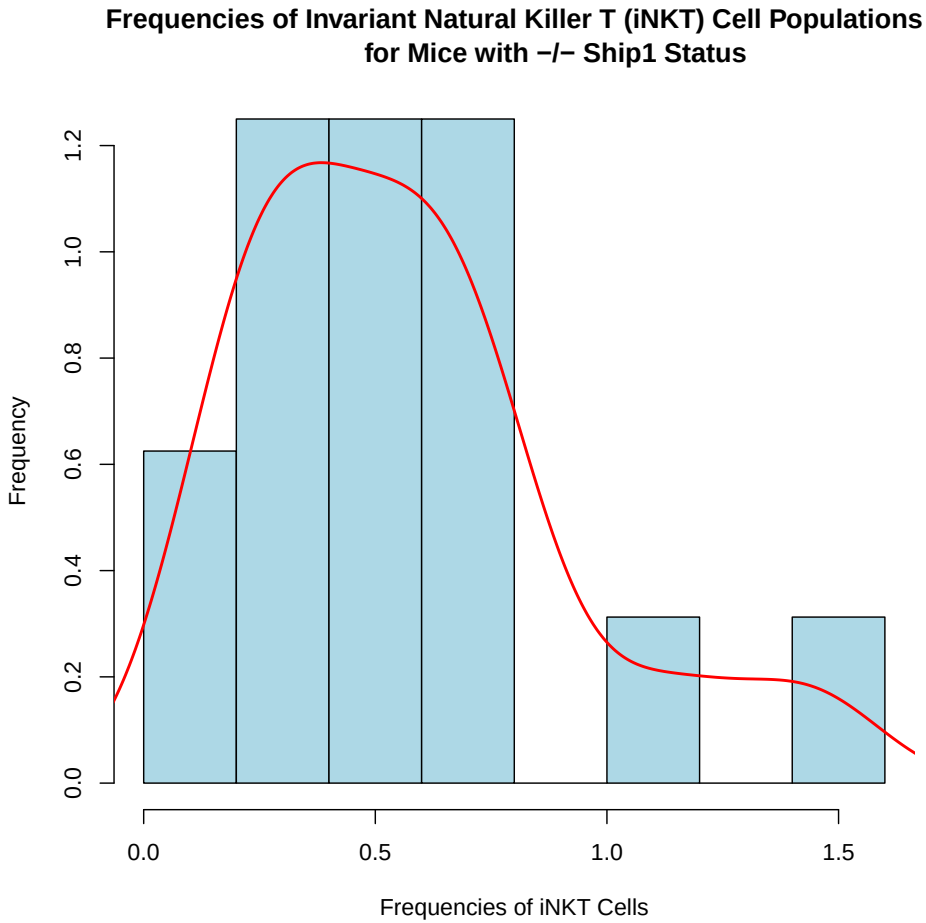
Graphical Summary – Histogram



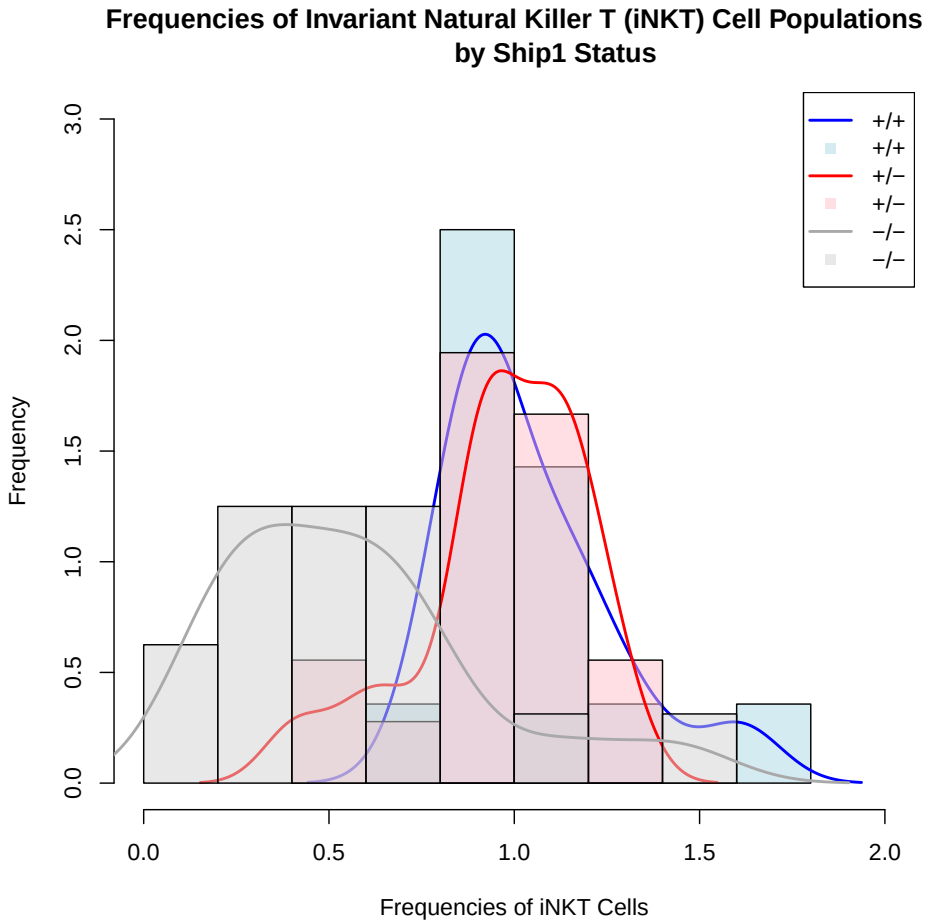
Graphical Summary – Histogram



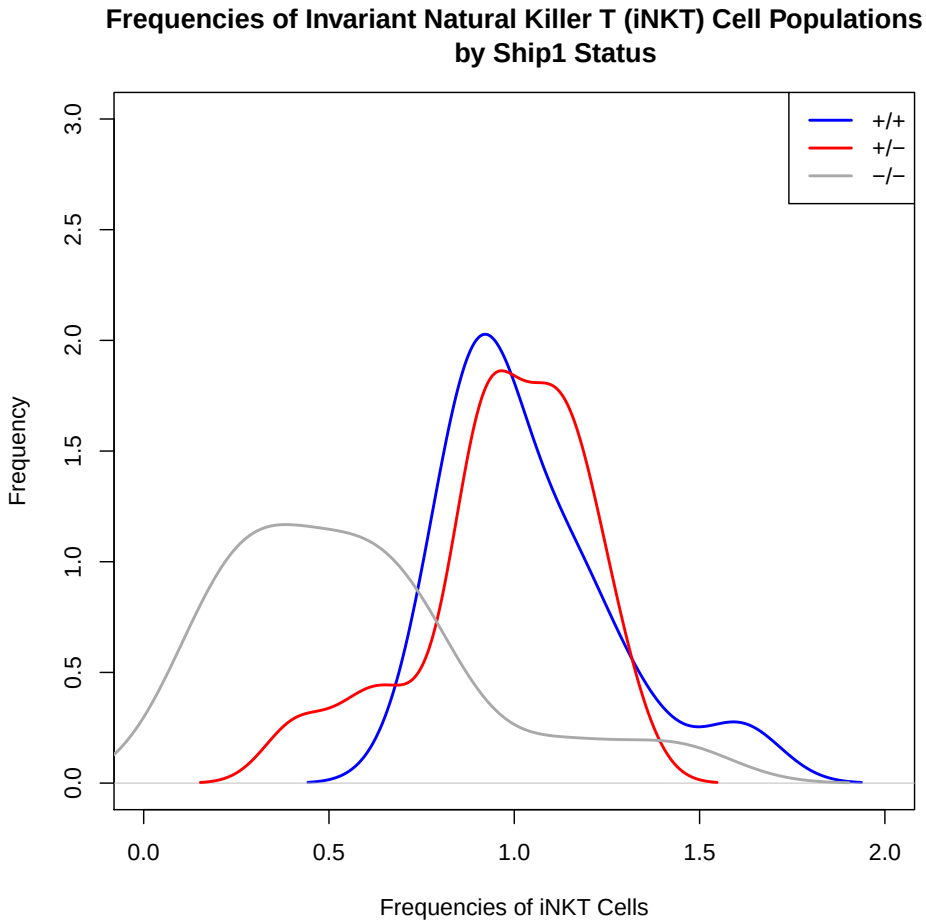
Graphical Summary – Histogram



Graphical Summary – Histogram



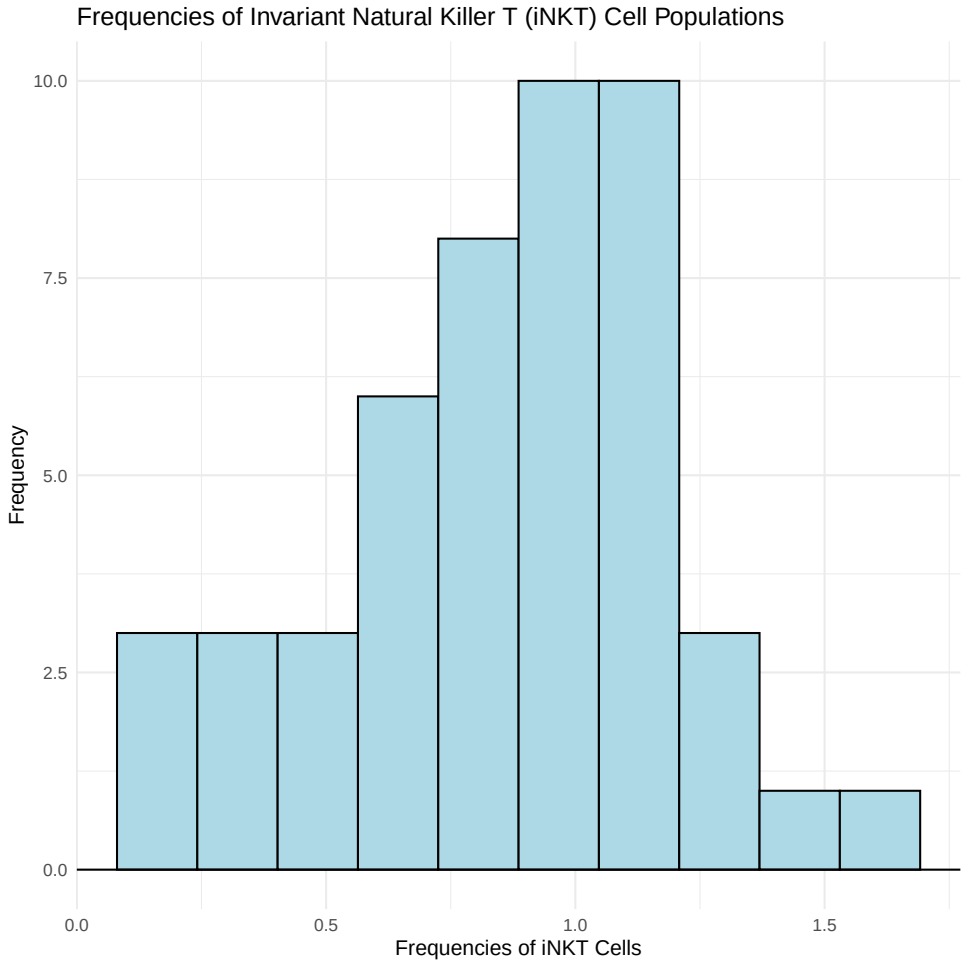
Graphical Summary – Histogram



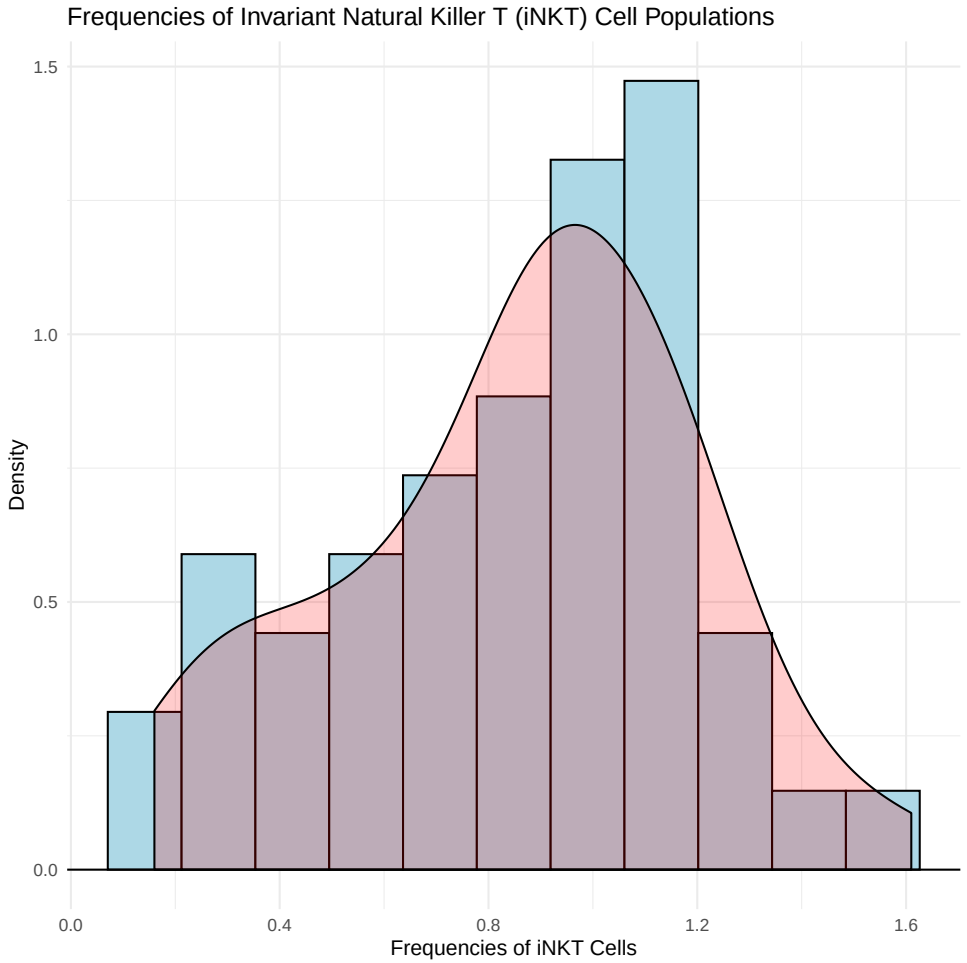
Graphical Summary – Histogram (ggplot)

Operator	Functionality
geom_histogram(...)	creates a histogram in ggplot
geom_density(...)	computes and draws kernel density estimate in gg

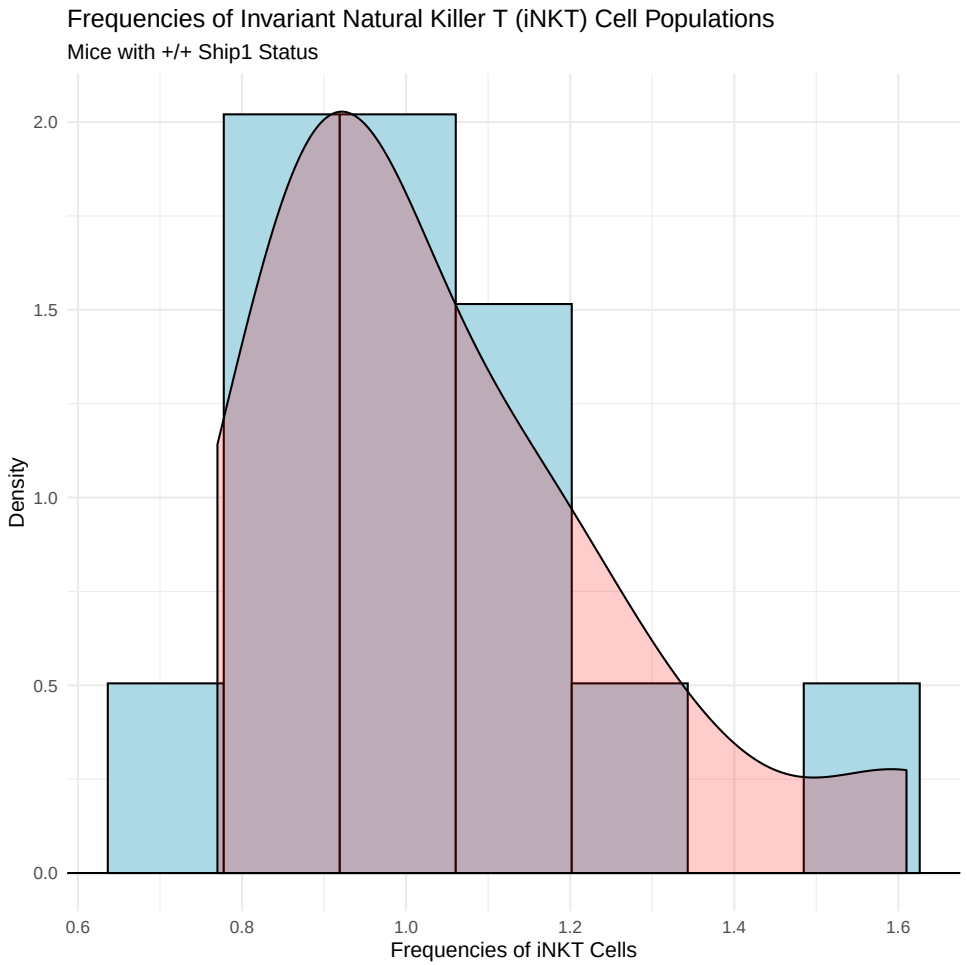
Graphical Summary – Histogram



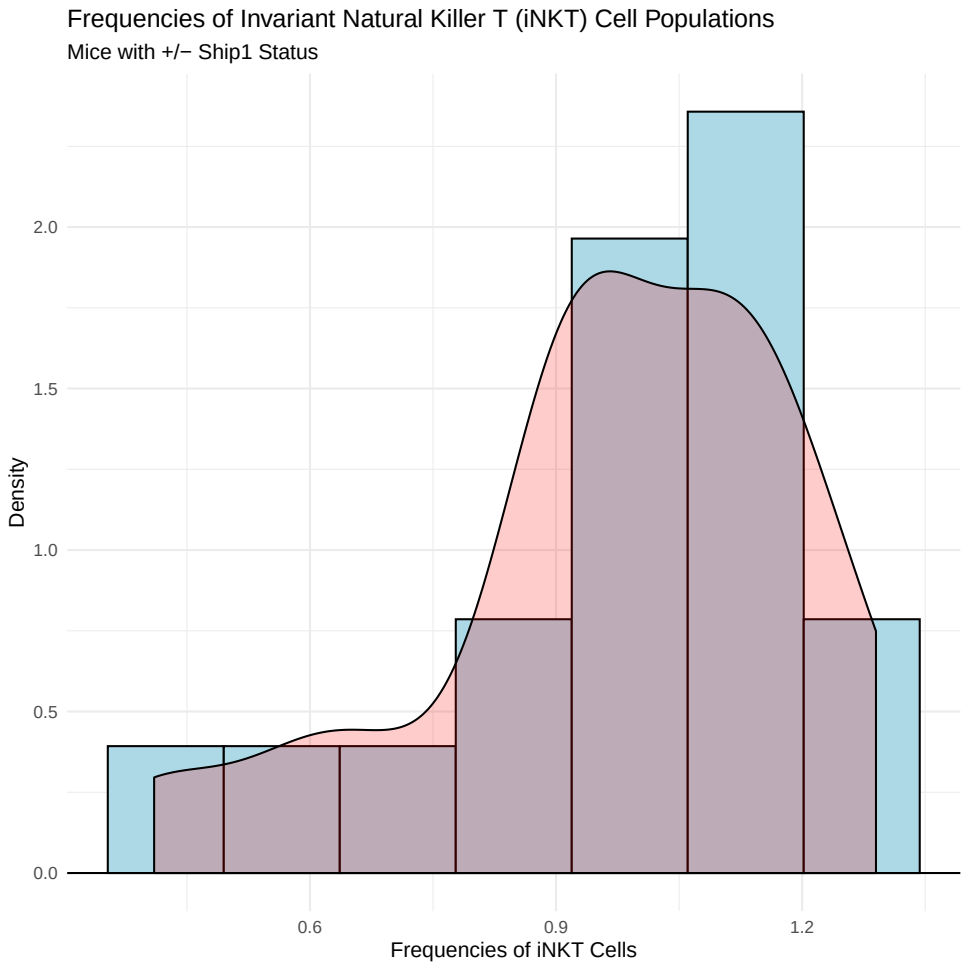
Graphical Summary – Histogram



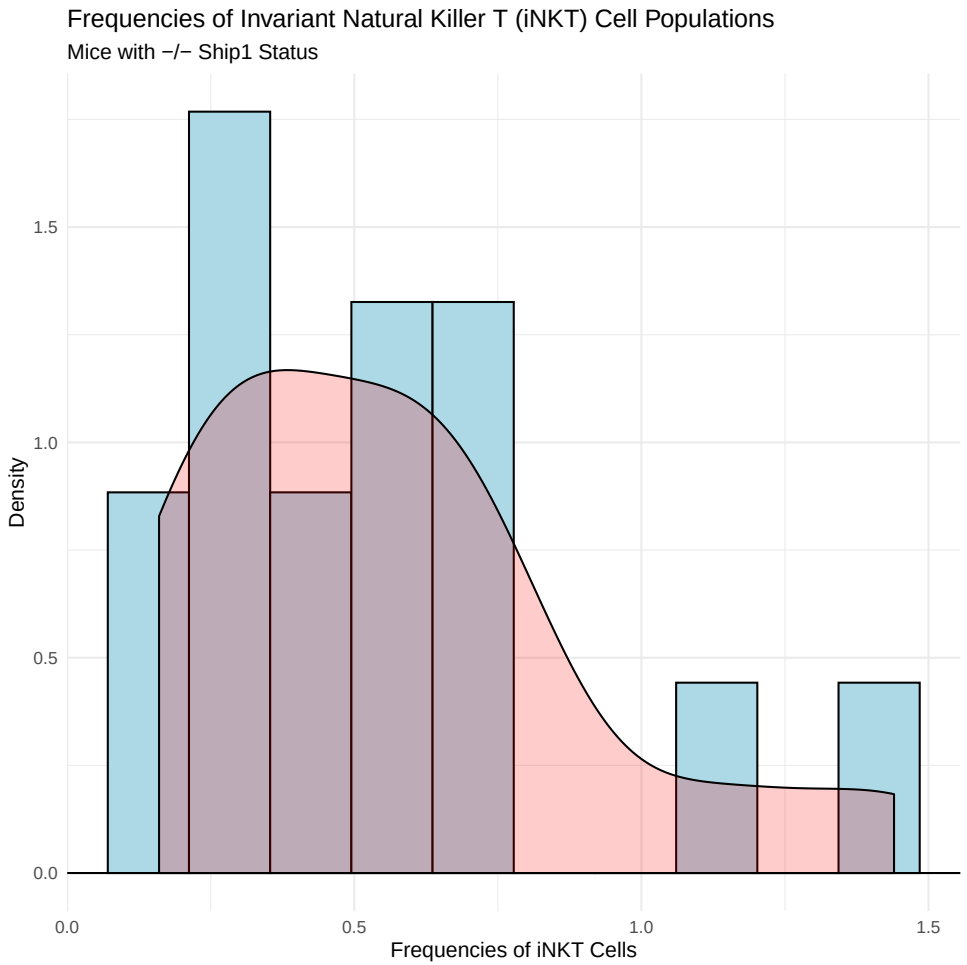
Graphical Summary – Histogram



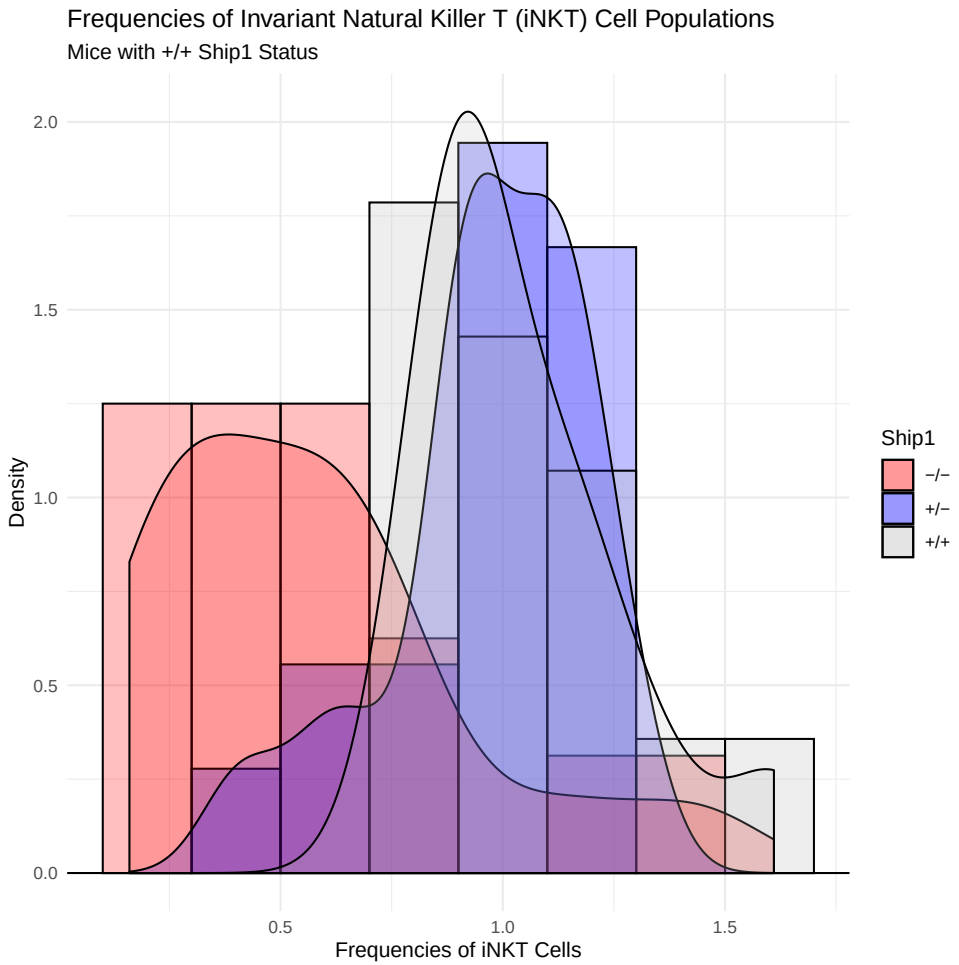
Graphical Summary – Histogram



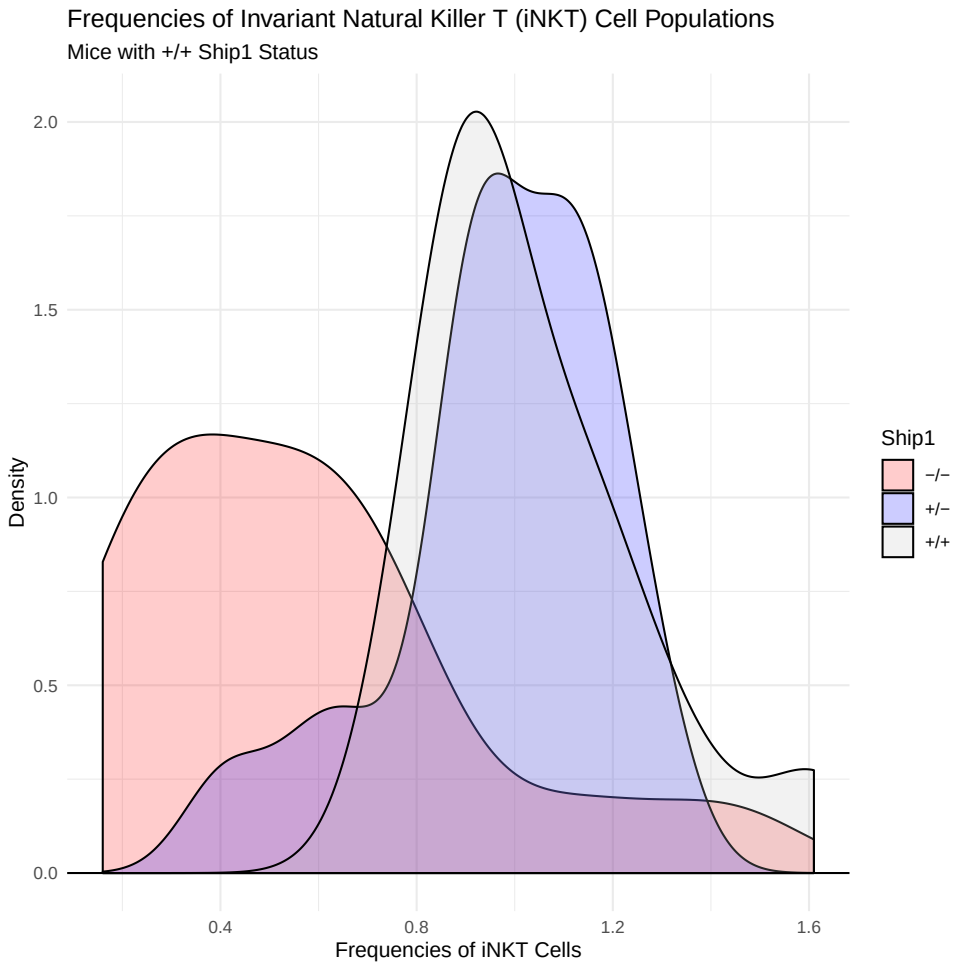
Graphical Summary – Histogram



Graphical Summary – Histogram



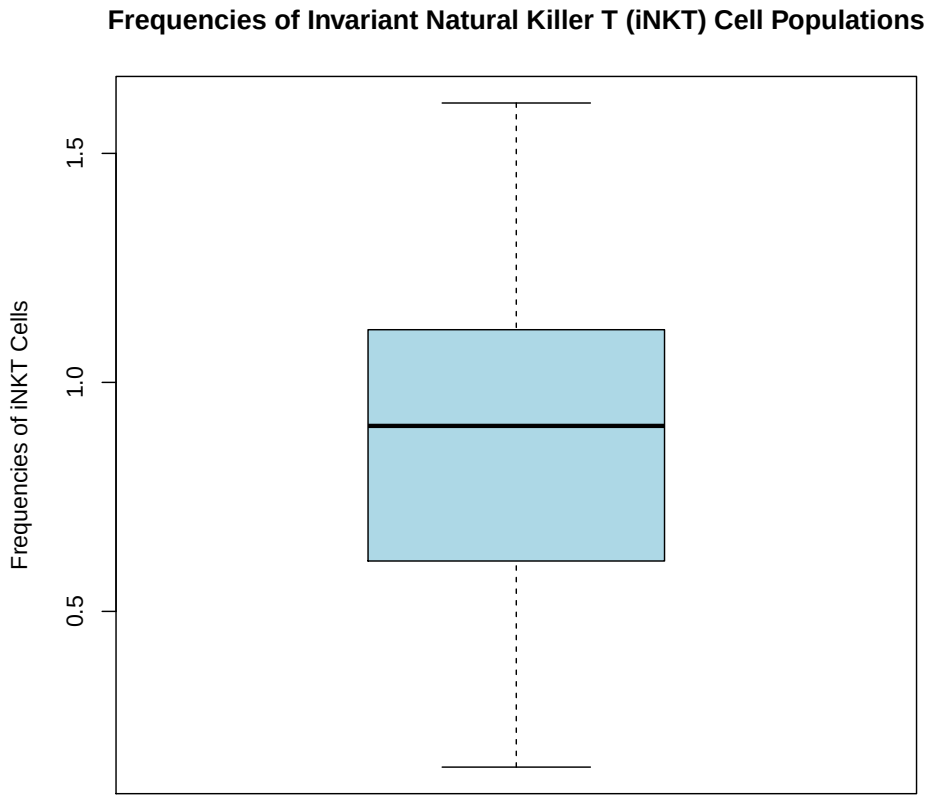
Graphical Summary – Histogram



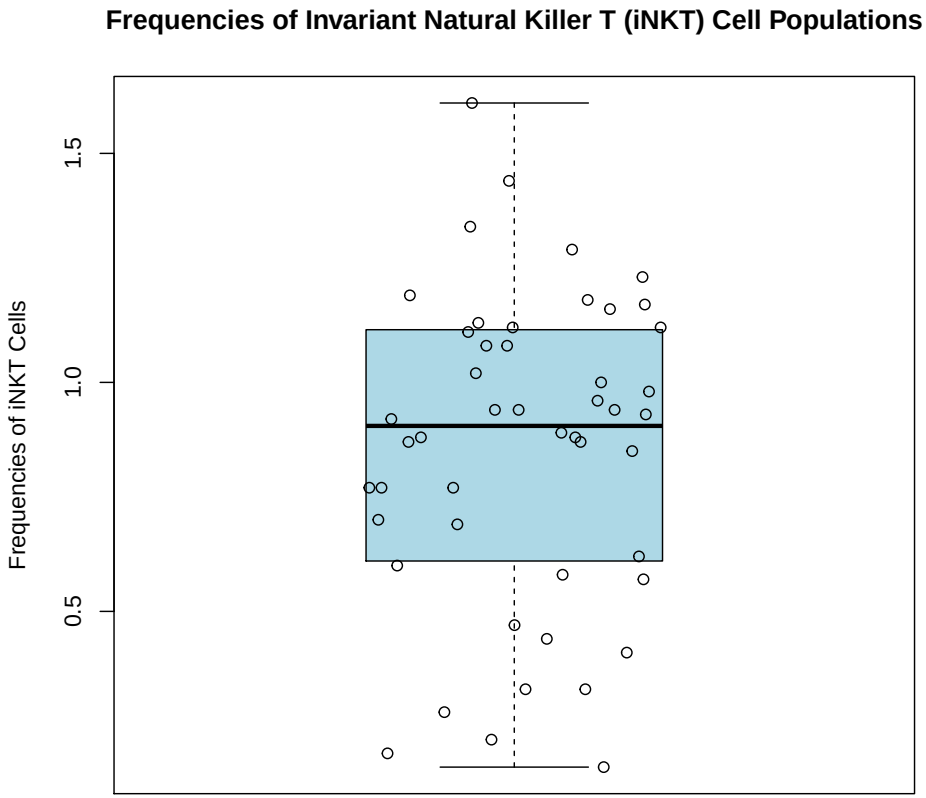
Graphical Summary – Boxplot

Operator	Functionality
boxplot(...)	creates a boxplot for data
points(...)	adds points to a plot
jitter(...)	adds noise to a vector

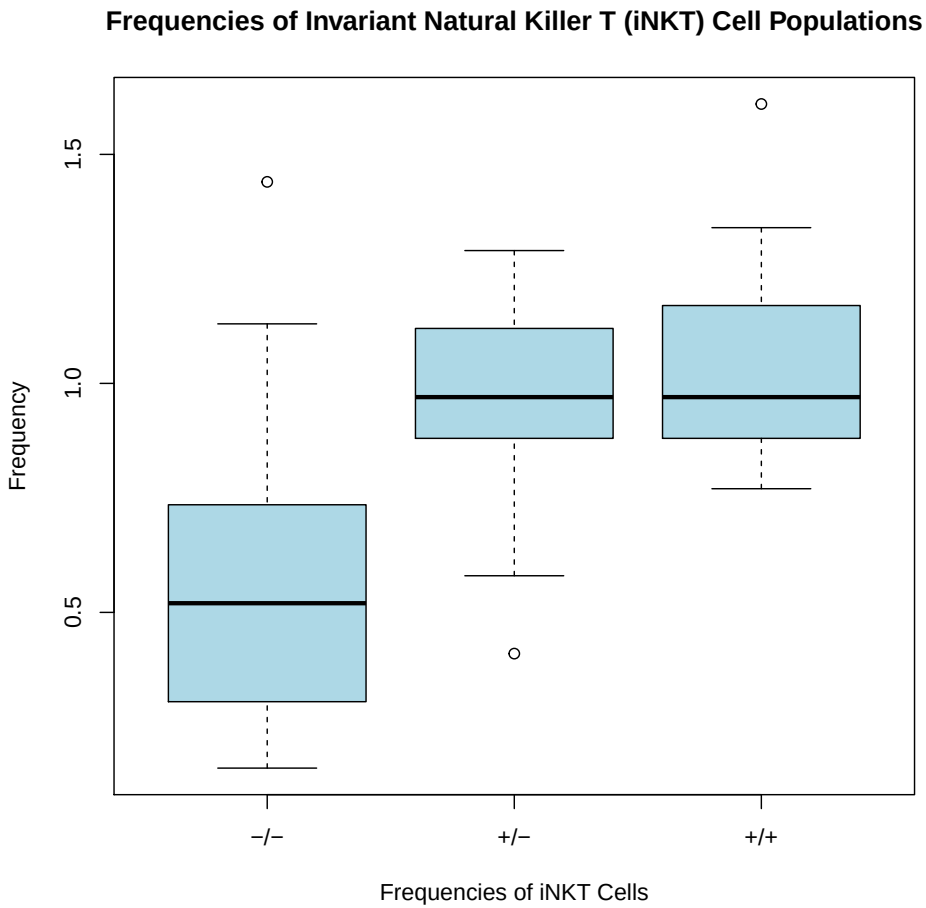
Graphical Summary – Boxplot



Graphical Summary – Boxplot



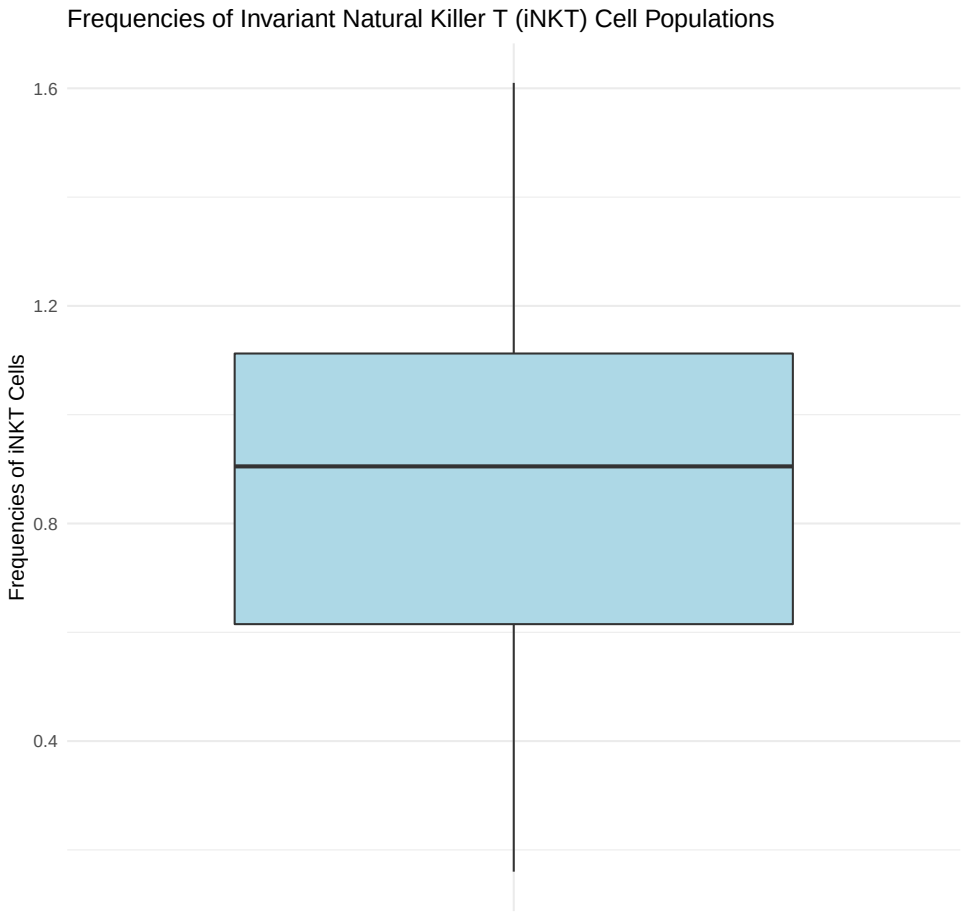
Graphical Summary – Boxplot



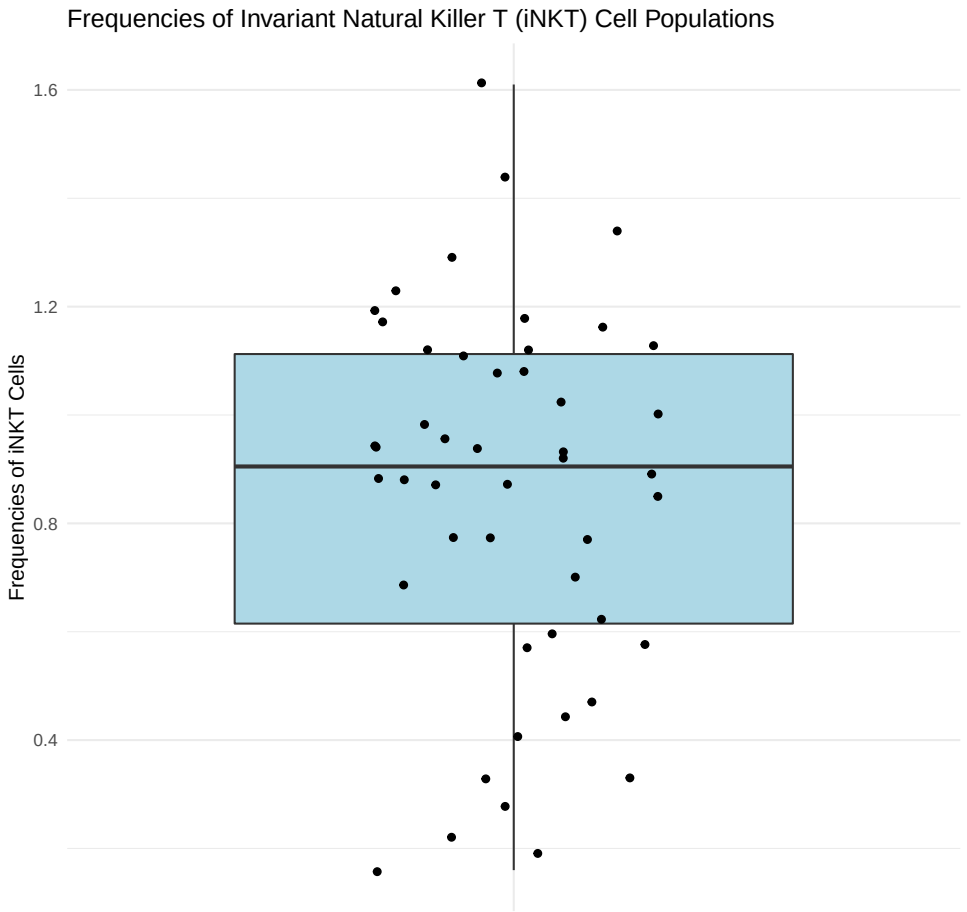
Graphical Summary – Boxplot (ggplot)

Operator	Functionality
geom_boxplot(...)	creates a boxplot in ggplot
geom_jitter(...)	adds noise and plots points in ggplot

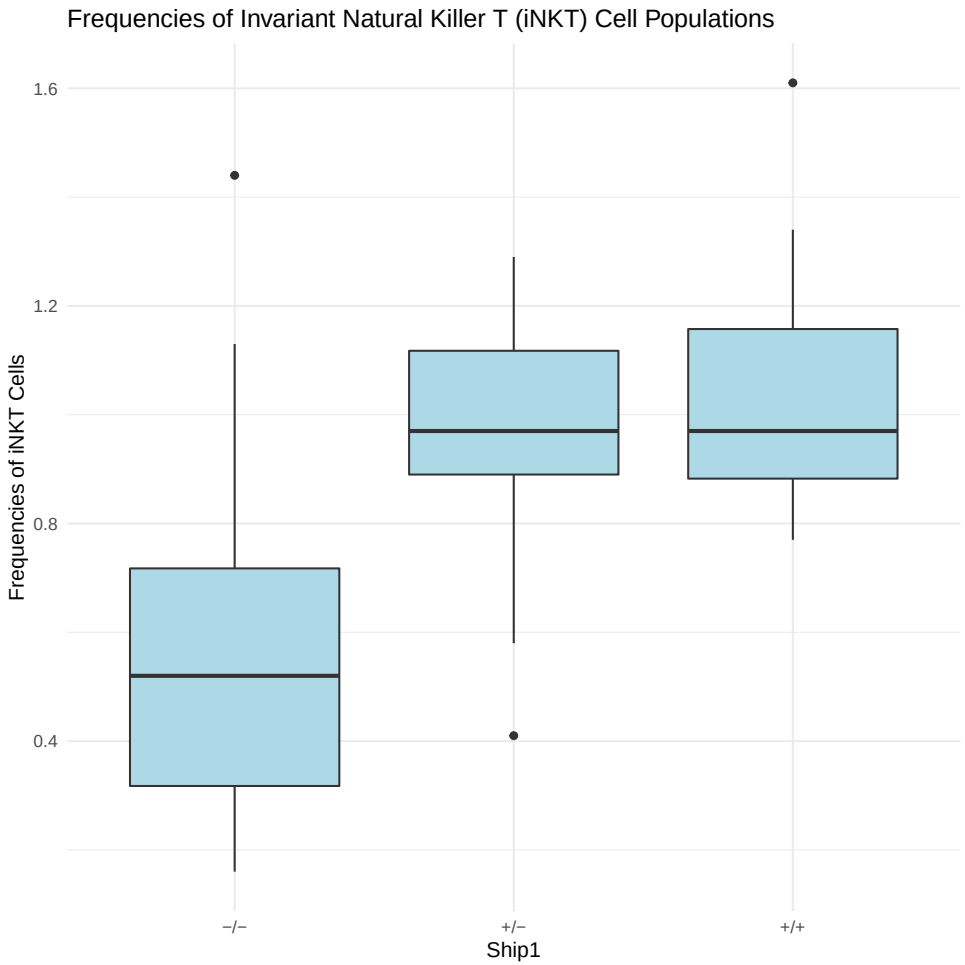
Graphical Summary – Boxplot



Graphical Summary – Boxplot



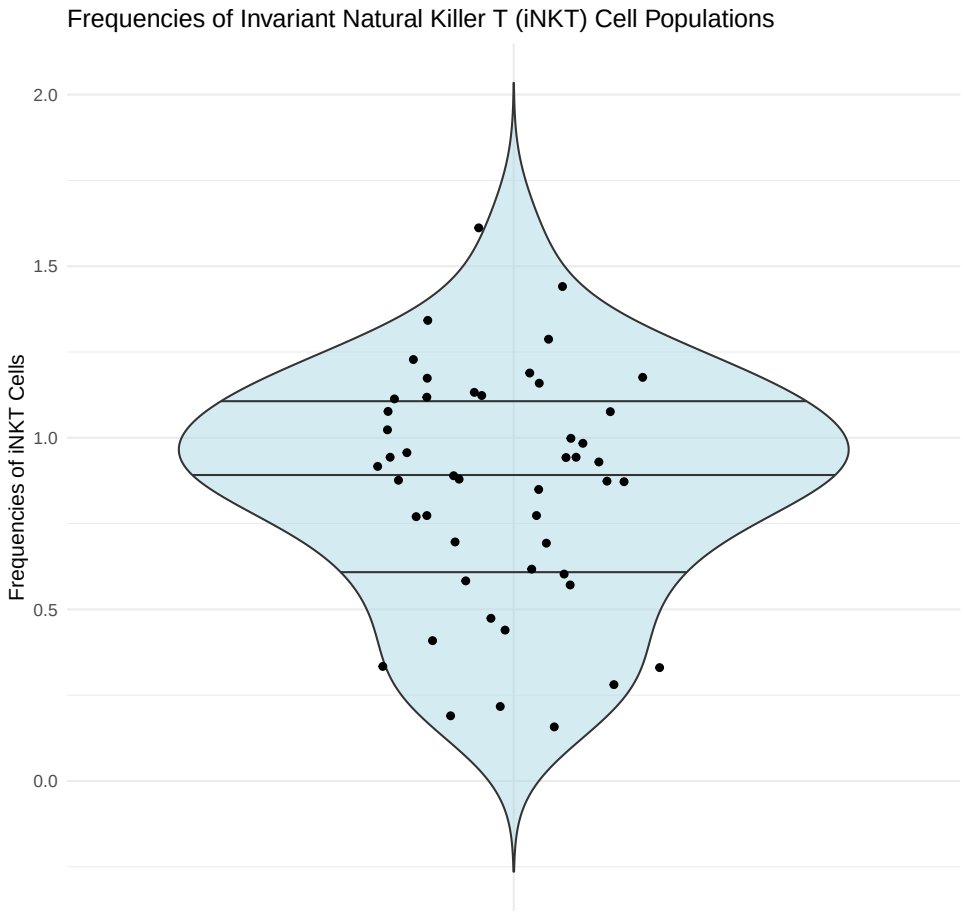
Graphical Summary – Boxplot



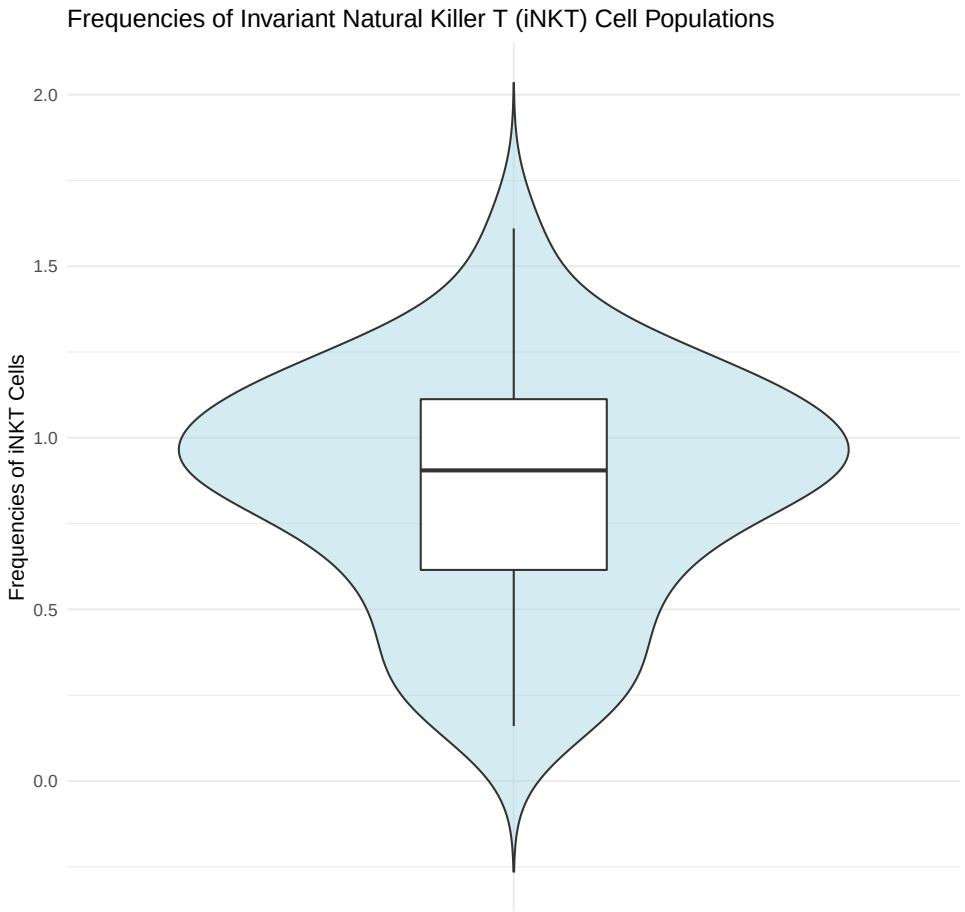
Graphical Summary – Violin Plot (ggplot)

Operator	Functionality
geom_violin(...)	creates a violin plot in ggplot

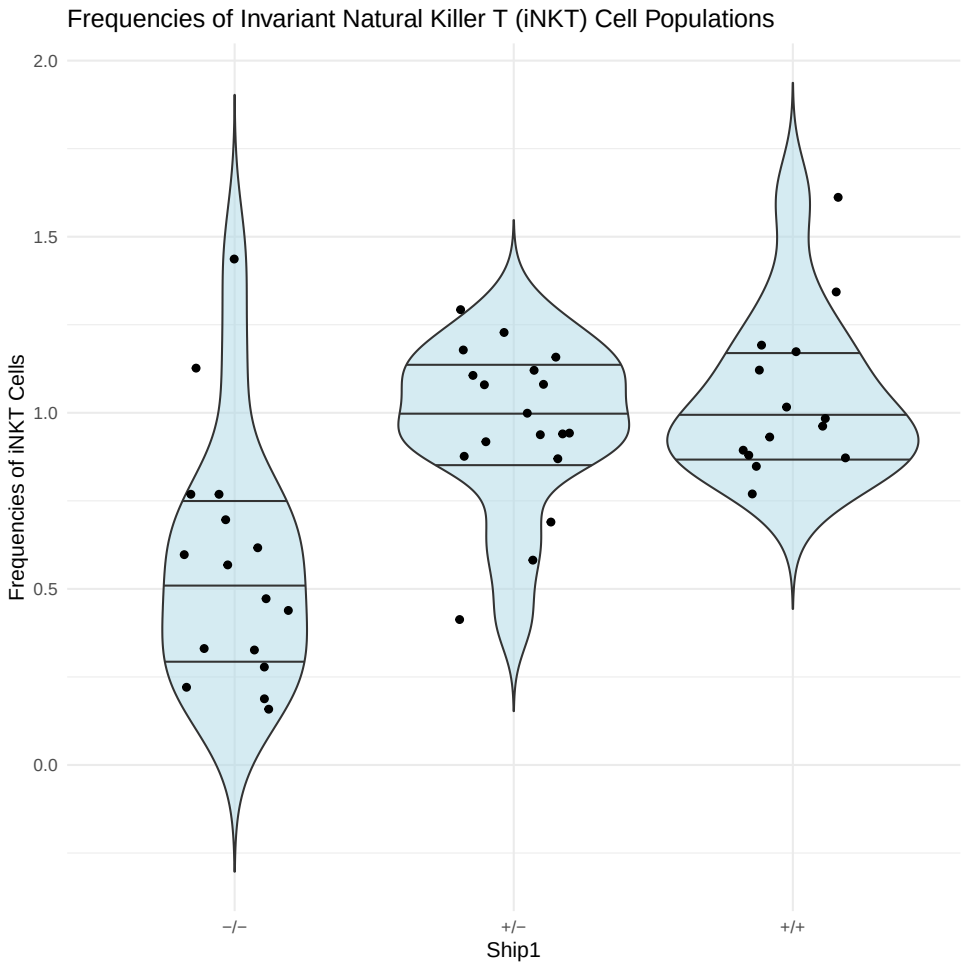
Graphical Summary – Violin Plot



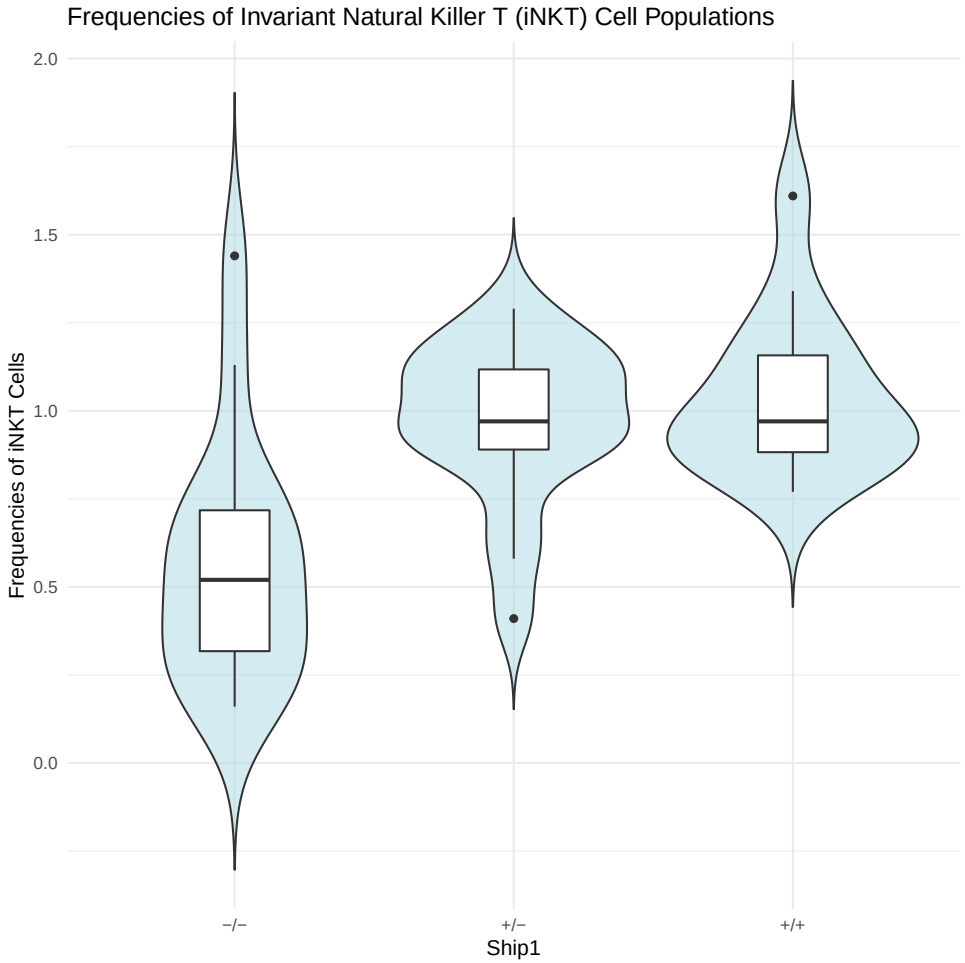
Graphical Summary – Violin Plot



Graphical Summary – Violin Plot



Graphical Summary – Violin Plot



Next Time

- More Case Study
- More Plotting
- Probability Distributions