Assignment 2—One-way Analysis of Variance.

Yudhajit Ain

Student ID: 30182745

University of Calgary

PSYC 615—Lab

TAs: Christopher Davie & Benjamin Moon

Assignment 2—One-way Analysis of Variance.

# Question 1

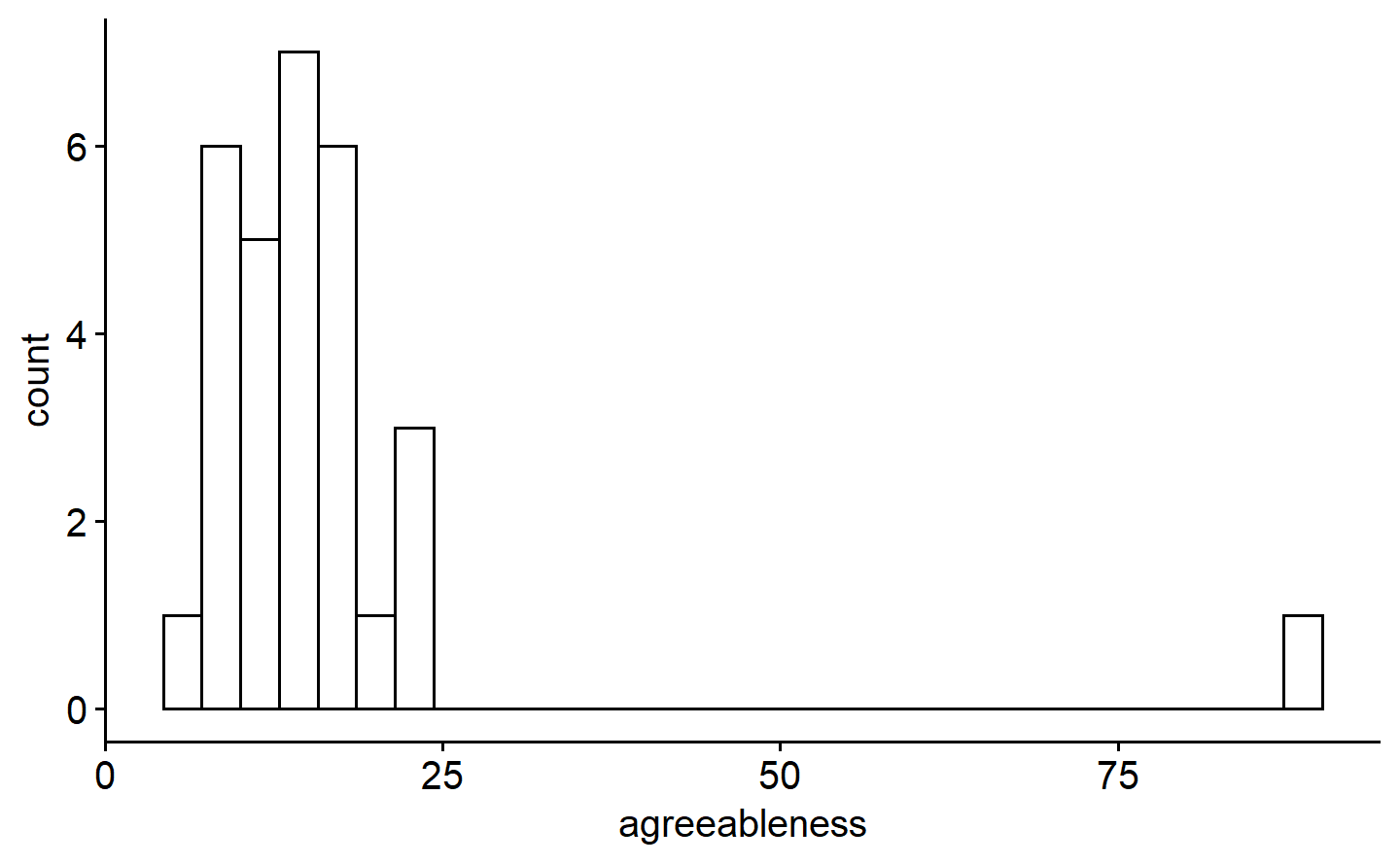
*From the data directory, please import assignment-2\_q1.csv into R.*

*Background: Nick Fury has approached you. He wants to put together a band of superheroes who can save the universe from current and future supervillains. He needs to decide whether he should hire superheroes from the Marvel universe (1), DC comics (2), Disney Princesses (3), or Smurfs (4). Because he plans to randomly sneak up on them to invite them (rather than, you know, sending an email) he wants to choose the group with highest agreeableness scores on the HEXACO personality inventory to avoid nasty confrontations when he shows up unannounced. He brings a dataset of agreeableness scores (on a scale from 1 – 34) to you and asks you to tell him if one group has significantly higher agreeableness, and which group that is.*

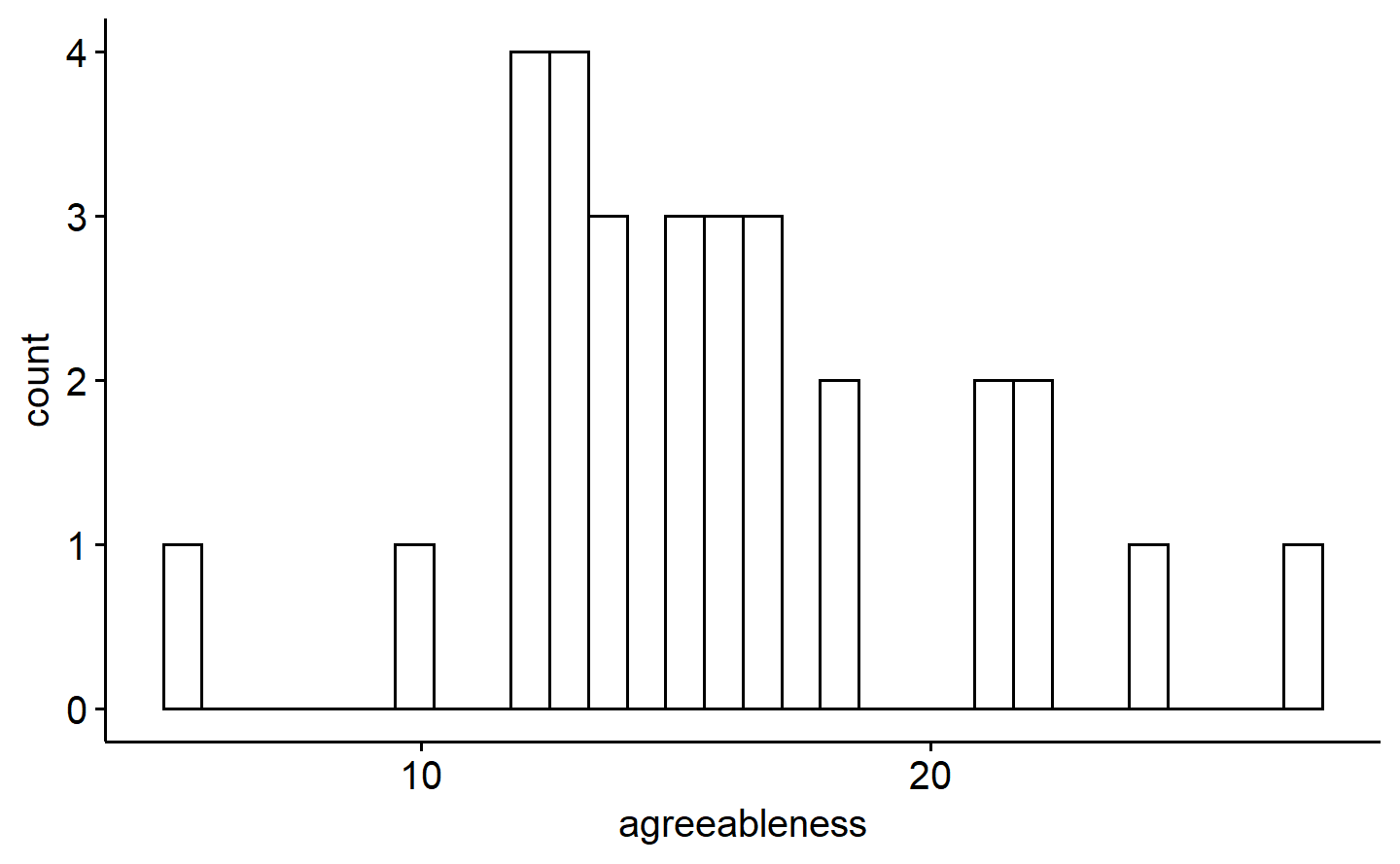
## Question 1a

*Let’s start with exploring the data. Does your data meet the assumptions for an ANOVA? If not, describe the assumption(s) that are violated, decide how you want to address the violations and explain exactly what you did, then re-check and report on your assumptions. (4 marks) Hint: Remember ANOVA is robust to violations of assumptions only in some circumstances – are those circumstances met?*

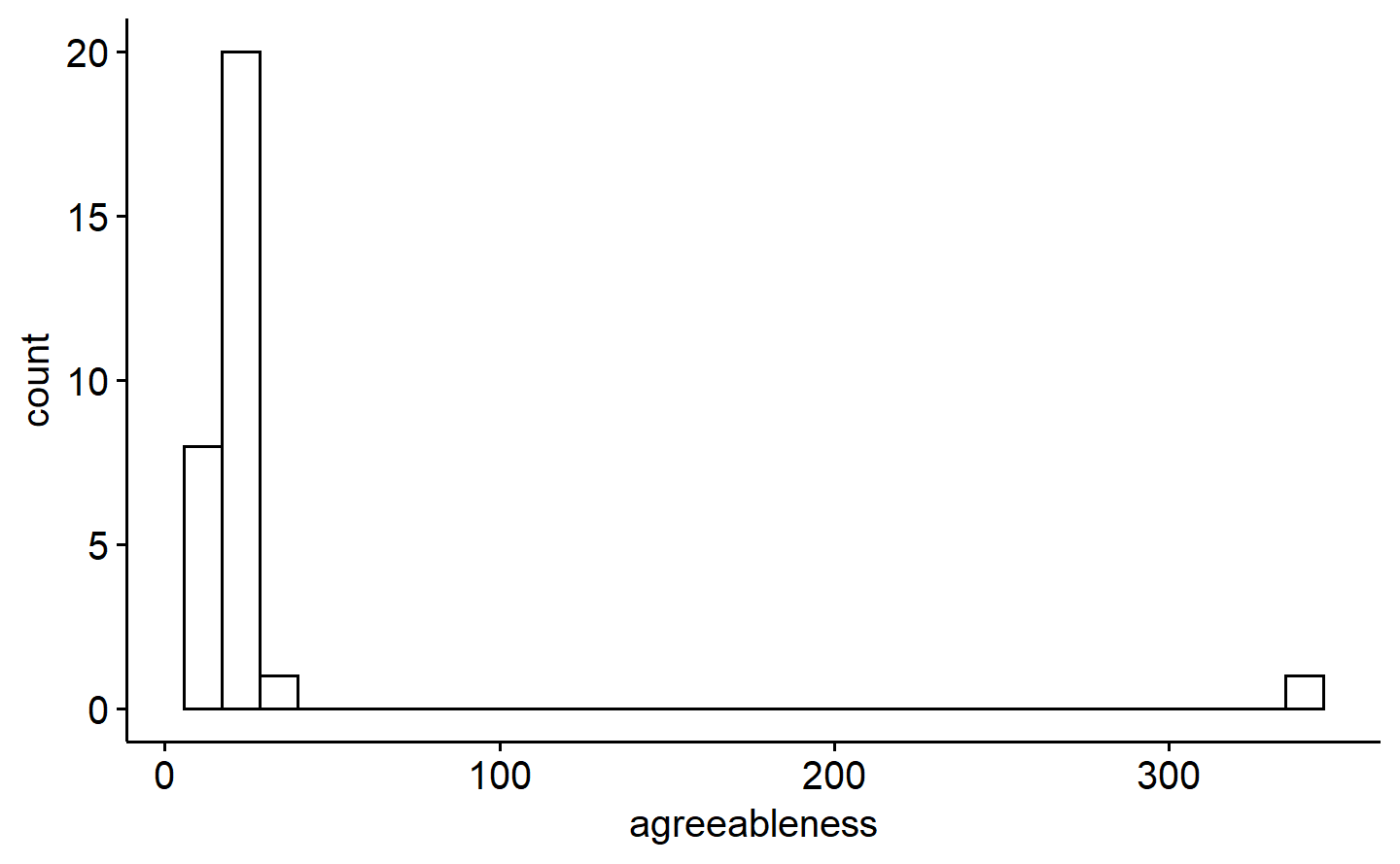
#>   
#> Shapiro-Wilk normality test  
#>   
#> data: data\_q1\_dc$agreeableness  
#> W = 0.44364, p-value = 1.427e-09  
#>   
#> Shapiro-Wilk normality test  
#>   
#> data: data\_q1\_marvel$agreeableness  
#> W = 0.9566, p-value = 0.2529  
#>   
#> Shapiro-Wilk normality test  
#>   
#> data: data\_q1\_smurfs$agreeableness  
#> W = 0.24278, p-value = 2.397e-11  
#>   
#> Shapiro-Wilk normality test  
#>   
#> data: data\_q1\_disney$agreeableness  
#> W = 0.95001, p-value = 0.1692



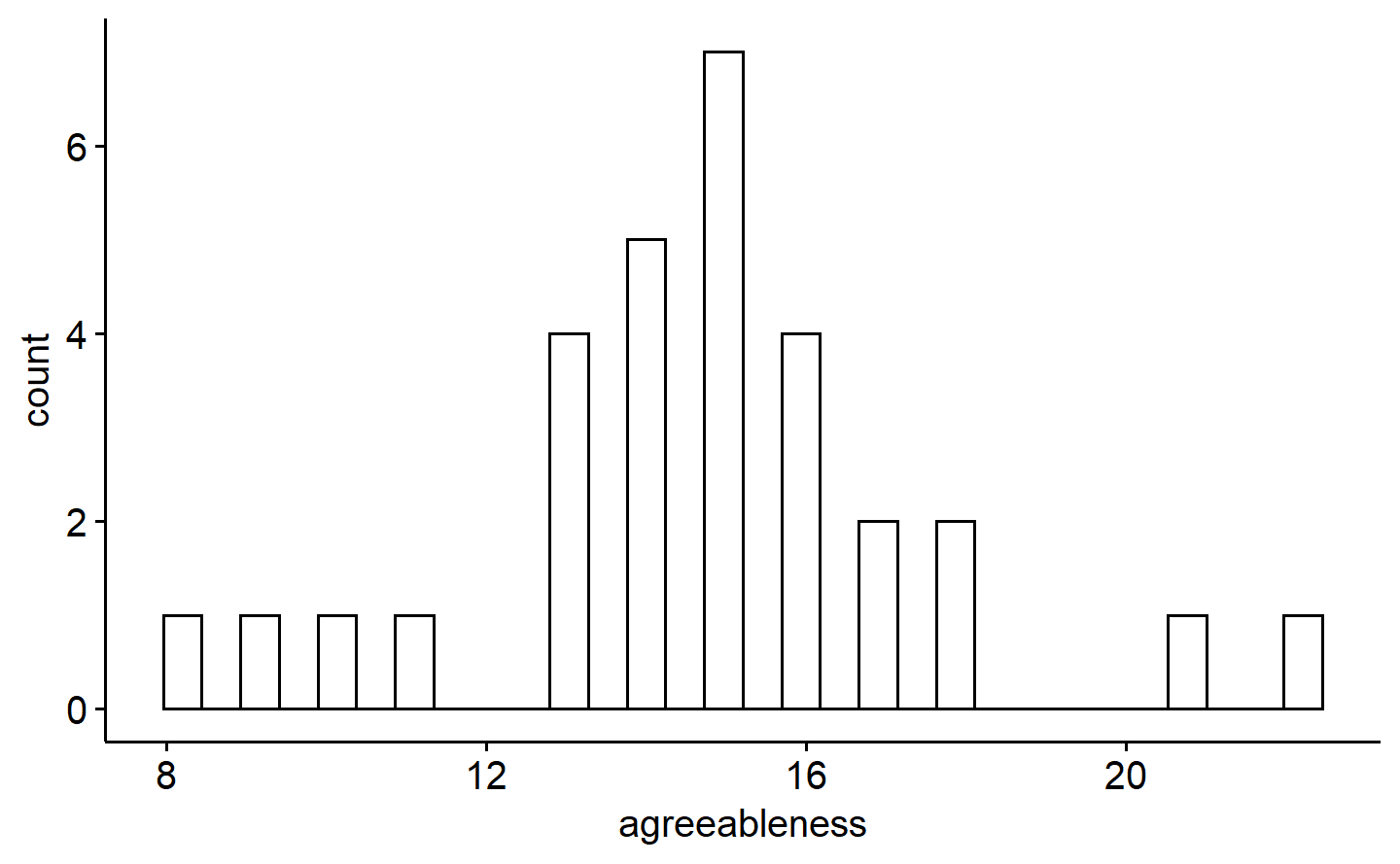
*Figure* *1.*  Distribution of agreeableness for DC superheroes



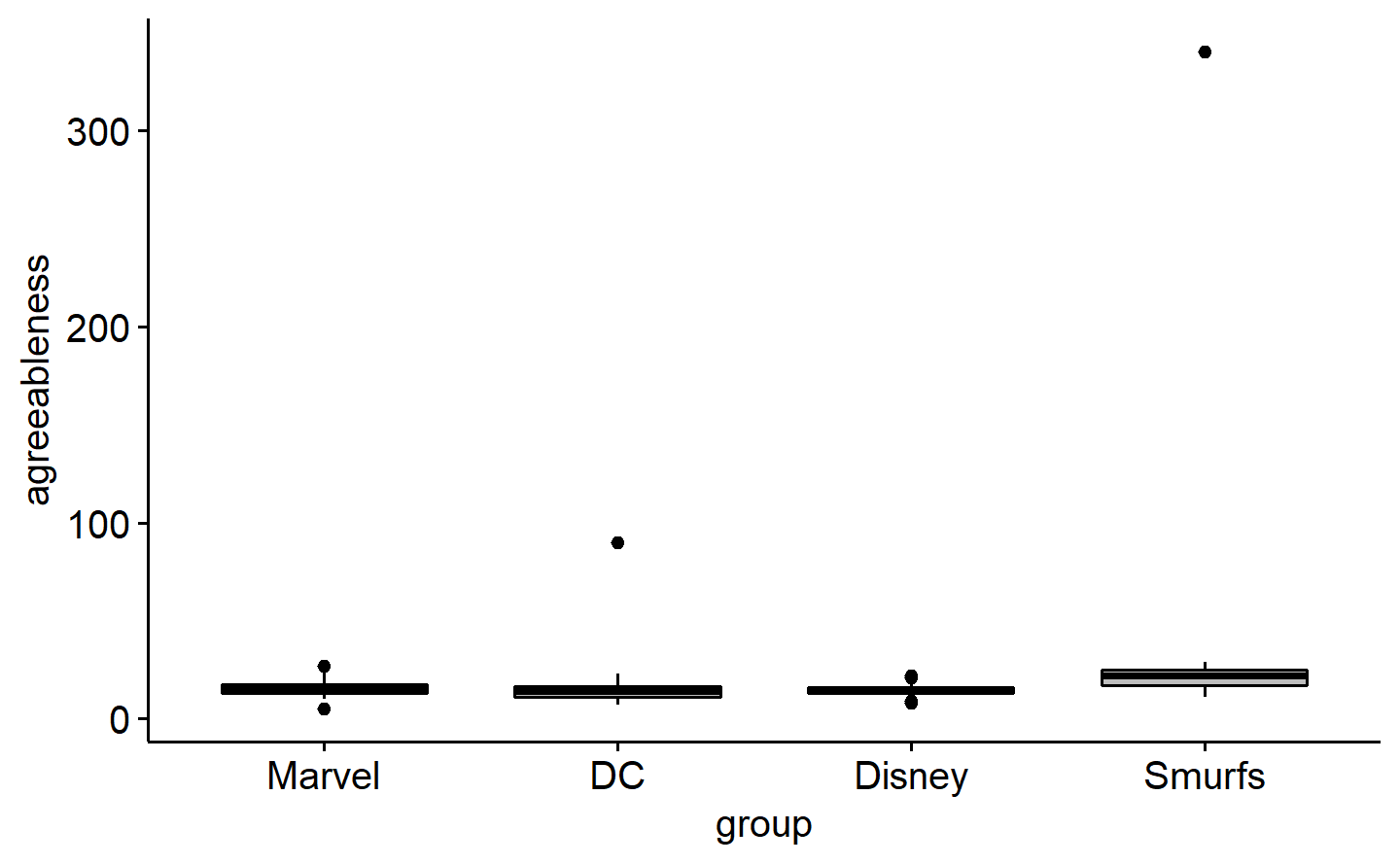
*Figure* *2.*  Distribution of agreeableness for Marvel superheroes



*Figure* *3.*  Distribution of agreeableness for Smurfs



*Figure* *4.*  Distribution of agreeableness for Disney princesses



*Figure* *5.*  Box-plot of agreeableness for all groups

#> # A tibble: 120 × 3  
#> superhero group agreeableness  
#> <dbl> <chr> <dbl>  
#> 1 21 Marvel 5  
#> 2 118 DC 7  
#> 3 14 DC 8  
#> 4 27 Disney 8  
#> 5 66 DC 8  
#> 6 106 DC 8  
#> 7 107 Disney 9  
#> 8 114 DC 9  
#> 9 30 DC 10  
#> 10 38 DC 10  
#> # … with 110 more rows  
#> Levene's Test for Homogeneity of Variance (center = mean)  
#> Df F value Pr(>F)   
#> group 3 2.8296 0.04162 \*  
#> 114   
#> ---  
#> Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
#>   
#> Shapiro-Wilk normality test  
#>   
#> data: data\_q1\_filtered\_dc$agreeableness  
#> W = 0.96266, p-value = 0.3815  
#>   
#> Shapiro-Wilk normality test  
#>   
#> data: data\_q1\_filtered\_marvel$agreeableness  
#> W = 0.9566, p-value = 0.2529  
#>   
#> Shapiro-Wilk normality test  
#>   
#> data: data\_q1\_filtered\_smurfs$agreeableness  
#> W = 0.96108, p-value = 0.3493  
#>   
#> Shapiro-Wilk normality test  
#>   
#> data: data\_q1\_filtered\_disney$agreeableness  
#> W = 0.95001, p-value = 0.1692  
#> [1] 19.13793  
#> [1] 20.30345  
#> [1] 21.83744  
#> [1] 8.96092

Answers here.

We plotted separtae histograms to look at the distribution of agreeableness in all superhero groups (figures 1-4), and also explicitly tested normality within groups using the Shapiro-Wilk test. The distribution was significantly non-normal for both the DC and Smurfs group.

Next, we made a box-plot of the agreeableness score for all groups (figure 5), to look at possible outliers. Both non-normal groups (DC and Smurfs) showed positive outliers, and they were then filtered out.

After filtering, the Shapiro-Wilk normality test was repeated, and all groups had a normal distribution now. Thereafter, we did the Levene’s test to test for homogeneity of variance across groups, and it revealed that the variances were not equal across all groups. However, since 1) the distributions are now all normal, 2) the sample sizes are almost equal for all groups, and finally, 3) the highest group variance (19.14 for DC) is less than 3 times the lowest variance (8.96 for Disney), the ANOVA should be fairly robust to the violation of homogeneity of variance.

## Question 1b

*Evaluate if the superhero groups differ on agreeableness and, if so, which groups differ. Write an APA formatted results section. (10 marks) This includes stating the hypothesis that is being evaluated, quickly reporting on assumptions (we assume independence from here on so no need for that, and only report normality and homogeneity after removing any outliers), and reporting the findings, including the omnibus test and all follow-ups (provide and reference an APA formatted figure). Hint: When reporting findings, you need to include practical significance!*

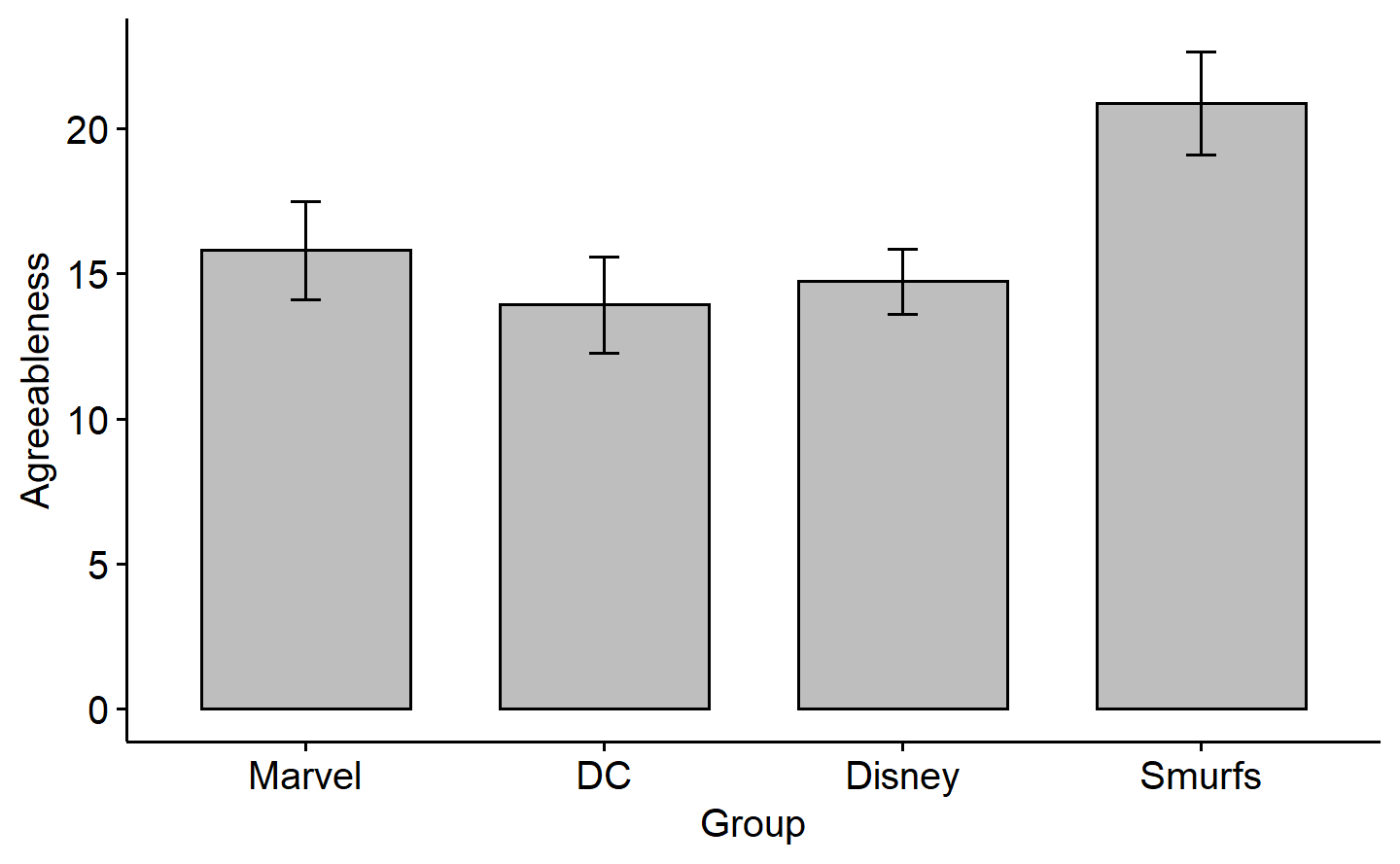
# Results

#> # A tibble: 2 × 6  
#> term df sumsq meansq statistic p.value  
#> <chr> <dbl> <dbl> <dbl> <dbl> <dbl>  
#> 1 group 3 847. 282. 16.1 8.30e-9  
#> 2 Residuals 114 1996. 17.5 NA NA  
#> Call:  
#> aov(formula = agreeableness ~ group, data = data\_q1\_filtered)  
#>   
#> Terms:  
#> group Residuals  
#> Sum of Squares 847.4213 1995.9770  
#> Deg. of Freedom 3 114  
#>   
#> Residual standard error: 4.184324  
#> Estimated effects may be unbalanced  
#> # Effect Size for ANOVA  
#>   
#> Parameter | Omega2 | 95% CI  
#> ---------------------------------  
#> group | 0.28 | [0.16, 1.00]  
#>   
#> - One-sided CIs: upper bound fixed at [1.00].  
#> group emmean SE df lower.CL upper.CL  
#> DC 13.9 0.777 114 12.4 15.5  
#> Disney 14.7 0.764 114 13.2 16.2  
#> Marvel 15.8 0.764 114 14.3 17.3  
#> Smurfs 20.9 0.777 114 19.3 22.4  
#>   
#> Confidence level used: 0.95  
#> contrast estimate SE df t.ratio p.value  
#> DC - Disney -0.802 1.09 114 -0.736 0.8823  
#> DC - Marvel -1.869 1.09 114 -1.715 0.3206  
#> DC - Smurfs -6.931 1.10 114 -6.307 <.0001  
#> Disney - Marvel -1.067 1.08 114 -0.987 0.7570  
#> Disney - Smurfs -6.129 1.09 114 -5.624 <.0001  
#> Marvel - Smurfs -5.062 1.09 114 -4.646 0.0001  
#>   
#> P value adjustment: tukey method for comparing a family of 4 estimates  
#> contrast effect.size SE df lower.CL upper.CL  
#> DC - Disney -0.192 0.261 114 -0.708 0.3248  
#> DC - Marvel -0.447 0.262 114 -0.966 0.0725  
#> DC - Smurfs -1.656 0.285 114 -2.220 -1.0926  
#> Disney - Marvel -0.255 0.259 114 -0.768 0.2577  
#> Disney - Smurfs -1.465 0.278 114 -2.015 -0.9142  
#> Marvel - Smurfs -1.210 0.272 114 -1.750 -0.6700  
#>   
#> sigma used for effect sizes: 4.184   
#> Confidence level used: 0.95  
#> [1] 13.93103  
#> [1] 15.8  
#> [1] 20.86207  
#> [1] 14.73333  
#> [1] 4.374692  
#> [1] 4.505935  
#> [1] 4.673055  
#> [1] 2.99348

Answers here.

# Results

In order to test for any significant mean differences on agreeableness between the four groups (DC, Marvel, Disney and Smurfs), we perform a one-way ANOVA with the null hypothesis : . Agreeableness was normally distributed in all groups, DC , Marvel , Smurfs and Disney princesses . The Levene’s test for homogeneity of variance, however, was significant , indicating unequal variance across the groups. With roughly equal sample sizes, and normality preserved within groups, we expect the ANOVA to be robust to the violation of homogeneity of variance, and so we proceed with it. The omnibus -test turns out to be significant , suggesting that at least one group has a significantly different mean agreeableness compared to other(s). To check for which specific groups have mean agreeableness significantly different from each other, we run a full pairwise post-hoc test, with -values adjusted using Tukey’s HSD. The results suggest that Smurfs were significantly more agreeable than DC heroes , , Marvel heroes , , as well as Disney princesses , . A bar-plot (figure 6) summarises the results visually.



*Figure* *6.*  Bar-plot of agreeableness across all groups

## Question 1c

*When Nick Fury reads your results, he gets very excited—he thinks this research could be written up in the Quarterly Journal of Superheroes and Government Agencies! He notices that one group appears to be significantly different from all the other groups and suggests conducting a complex comparison instead of pairwise comparisons as follow-up tests. State how you would respond to Nick Fury. Is it wise to run a complex comparison as a follow-up in this case? Why or why not? (2 marks)*

Answers here.

I would suggest Nick Fury to not go ahead with the idea. Having already done, and seen the results of the pairwise comparisons, the chances of making a type-I error in the complex comparison will be highly inflated because of the hypothesis being driven by results which are already significant. If he is very keen, however, I would suggest him to use a suitable test for this, like the -Scheffe’s.

# Question 2

## Question 2a

*Briefly describe how planned comparisons differ from post-hoc comparisons in terms of error? (1 mark)*

Answers here.

Planned comparisons are set up by the researcher(s) before having looked at the data, and is driven purely by the hypothesis and the theoretical predictions. Post-hoc comparisons, on the other hand, are done after having looked at the data visually, or after performing an omnibus -test or ANOVA. As such, the chances of making a type-I error is much more likely while doing uncorrected post-hoc comparisons (since the hypothesis is informed by the data itself), as opposed to uncorrected planned comparisons.

## Question 2b

*What are two different types of planned comparisons (think # of groups involved) a researcher can do and what makes them different? (1 mark)*

Answers here.

The two kind of planned comparisons are : 1) simple (where we compare one group against another single group in the dataset), and 2) complex (where we compare a single, or combination of groups against another combination of groups).

# Question 3

*From the data directory, please import assignment-2\_q3.csv into R.*

*Background: Did you know that people “research” sitcoms? For example, there are three main ’types” of sitcoms: House (surrounding a family home), Apartment (friends hanging out), and Workplace (co-worker shenanigans; see here!). Who knew, right? Now that you know this is a topic you can research, you are determined to see if different sitcoms are objectively “better” than others. You decide to recruit fellow graduate students, who are just as eager to find reasons to procrastinate, to watch three seasons of a sitcom and report on the following, very objective, measures: average laughs per episode, binge-worthiness of the show, and number of tasty snacks consumed while watching. These three variables will be summed to create one dependent variable “sitcom quality”. Trust it is valid & reliable!*

*Students are randomly assigned to watch one of the following sitcoms and note their measure values. The sitcoms used were:*

* *Modern family (House type)*
* *Friends (Apartment type)*
* *Seinfeld (Apartment type)*
* *The Office (Workplace type)*
* *Brooklyn Nine-Nine (Workplace type)*

*After your extensive background research in the area, which was definitely not just binging shows all weekend, you come up with the following three hypotheses (i.e., planned comparisons):*

1. *Home types (House & Apartment) are better than Workplace types.*
2. *Apartment types are better than House types.*
3. *Seinfeld is better than Friends.*

## Question 3a

*Are all of your planned comparisons orthogonal to each other? Please justify your answers by showing the formulas/calculations for each pair. (2 marks)*

Answers here.

The contrasts, with the “sitcom” factor in the same order (top to bottom) as in the “assignment-2\_q3.csv” dataset, are :

contrast 1 : (1,1,1,-1.5,-1.5)

contrast 2 : (-2,1,1,0,0)

constrat 3 : (0,-1,1,0,0)

For contrast\_1 and contrast\_2, the dot product =

For contrast\_1 and contrast\_3, the dot product =

For contrast\_2 and contrast\_3, the dot product =

Thus, all three comparisons are mutually orthogonal.

## Question 3b

*Explain why your planned comparisons are not a complete/full orthogonal set? State additional hypothesis(es)—in words and contrast codes—that could be used to complete the set. No need to show formulas/calculations for the new hypothesis. (1 mark)*

Answers here.

The number of groups (sitcoms) in our dataset = = 5. We know that the number of orthogonal contrasts that make up a complete orthogonal set for a model with groups, is . Here, that is four contrasts to make up a complete set. We already have three.

The last contrast that would complete the set is contrast 4 = (0,0,0,-1,1). This contrast essentially tests the hypothesis that Brooklyn Nine-Nine is better than The Office.

## Question 3c

*Write an APA results section that includes the following: (11 marks total, see below for breakdown):*

1. *Brief report of assumptions (No need to clean the data this time, aside from assumption checking. PLEASE DO NOT REMOVE ANY DATA FOR THIS ASSIGNMENT!) (1 mark)*
2. *Omnibus results (as is tradition, and including effect size) (1 mark)*
3. *The planned comparisons from Questions 3a and 3b (including means and SD that relate to the “groups” being compared and appropriate effect sizes) (8 marks)*
4. *Appropriate figures visualizing each hypothesis (3 marks).*

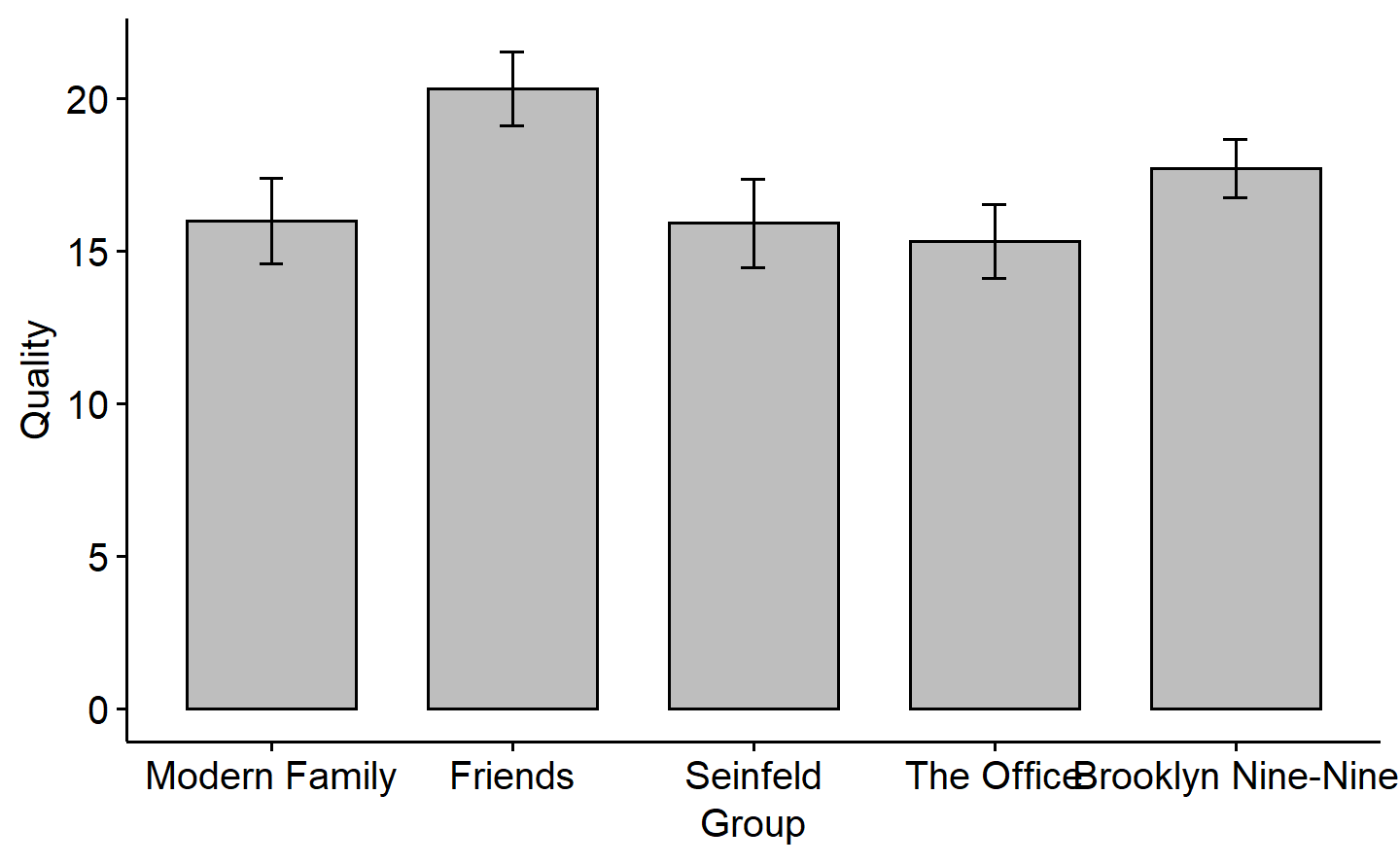
#>   
#> Shapiro-Wilk normality test  
#>   
#> data: data\_q3\_modfam$sitcom\_quality  
#> W = 0.94911, p-value = 0.2394  
#>   
#> Shapiro-Wilk normality test  
#>   
#> data: data\_q3\_friends$sitcom\_quality  
#> W = 0.9454, p-value = 0.197  
#>   
#> Shapiro-Wilk normality test  
#>   
#> data: data\_q3\_seinfeld$sitcom\_quality  
#> W = 0.94344, p-value = 0.1777  
#>   
#> Shapiro-Wilk normality test  
#>   
#> data: data\_q3\_office$sitcom\_quality  
#> W = 0.96992, p-value = 0.643  
#>   
#> Shapiro-Wilk normality test  
#>   
#> data: data\_q3\_b99$sitcom\_quality  
#> W = 0.94984, p-value = 0.2486  
#> Levene's Test for Homogeneity of Variance (center = mean)  
#> Df F value Pr(>F)  
#> group 4 1.5247 0.1993  
#> 120  
#> # A tibble: 2 × 6  
#> term df sumsq meansq statistic p.value  
#> <chr> <dbl> <dbl> <dbl> <dbl> <dbl>  
#> 1 sitcom 4 413. 103. 11.0 0.000000117  
#> 2 Residuals 120 1122. 9.35 NA NA  
#> # Effect Size for ANOVA  
#>   
#> Parameter | Omega2 | 95% CI  
#> ---------------------------------  
#> sitcom | 0.24 | [0.12, 1.00]  
#>   
#> - One-sided CIs: upper bound fixed at [1.00].  
#> # A tibble: 4 × 9  
#> term contrast null.…¹ estim…² std.e…³ df stati…⁴ adj.p…⁵ effec…⁶  
#> <chr> <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
#> 1 sitcom Home - Workplace 0 7.18 1.67 120 4.29 1.47e-4 0.783  
#> 2 sitcom Apartment - House 0 0.880 1.50 120 0.588 1 e+0 0.144  
#> 3 sitcom Seinfeld - Frien… 0 -4.32 0.865 120 -5.00 8.07e-6 -1.41   
#> 4 sitcom Brooklyn Nine-Ni… 0 -0.600 0.865 120 -0.694 1 e+0 -0.196  
#> # … with abbreviated variable names ¹​null.value, ²​estimate, ³​std.error,  
#> # ⁴​statistic, ⁵​adj.p.value, ⁶​effect.size  
#> [1] 17.41333  
#> [1] 3.866325  
#> [1] 16.52  
#> [1] 2.873098  
#> [1] 18.12  
#> [1] 3.910426  
#> [1] 16  
#> [1] 3.427827  
#> [1] 15.92  
#> [1] 3.534591  
#> [1] 20.32  
#> [1] 2.939955  
#> [1] 15.32  
#> [1] 2.939955  
#> [1] 17.72  
#> [1] 2.282542

Answers here.

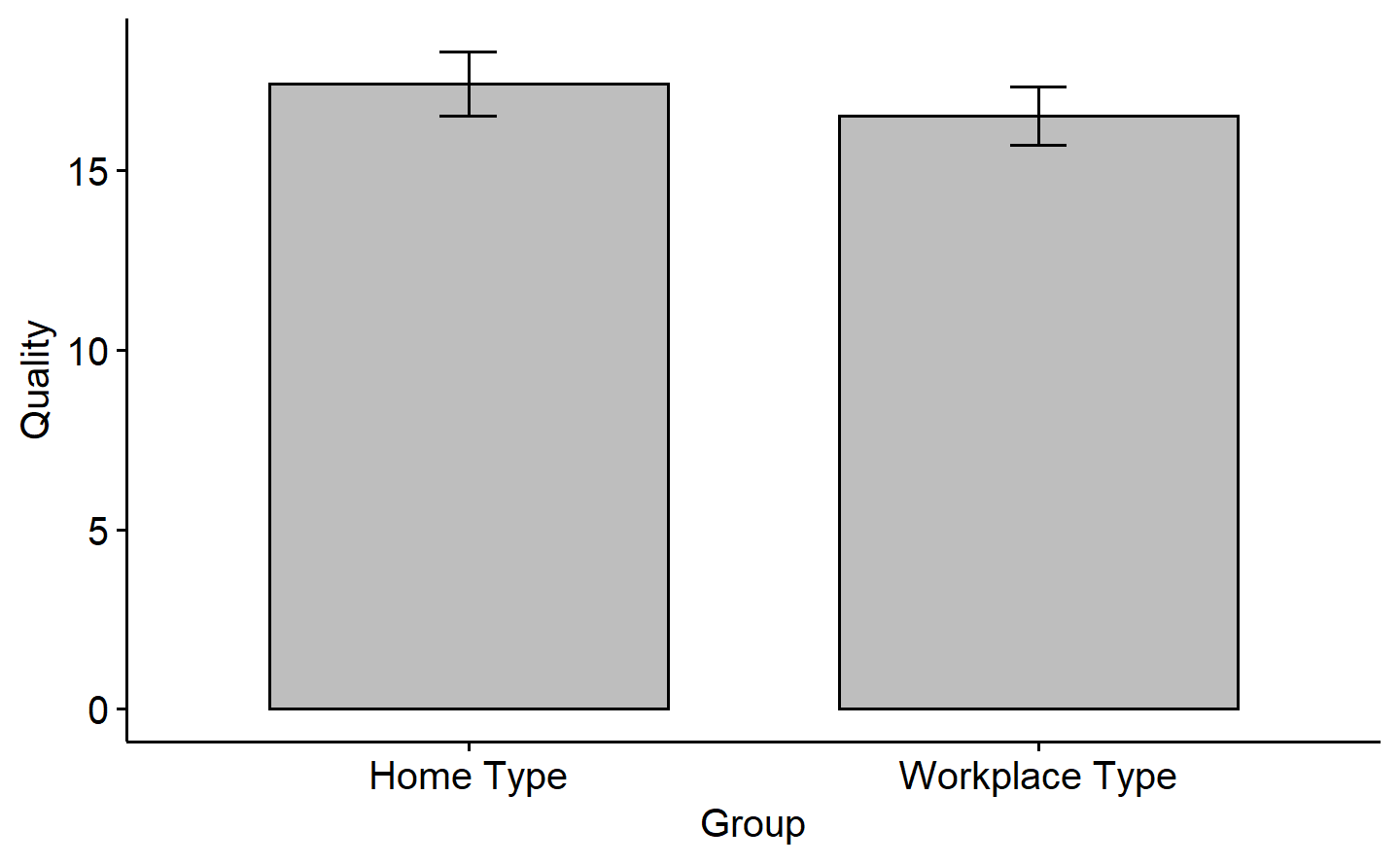
# Results

We wanted to test if there were any significant differences between the mean sitcom quality of the five sitcoms (groups) in our dataset. To do this, we have to first perform a one-way ANOVA. We check for normality within each group (sitcom), and find that the quality ratings for all of them, Modern Family , Friends , Seinfeld , The Office , and Brooklyn Nine-Nine were normally distributed. To check for homogeneity of variance across the groups, we perform the Levene’s Test, and the result suggests equality of variance across groups . With all assumptions met, we proceed to perform the one-way ANOVA. The results of the ANOVA suggest that indeed, at least one of the sitcoms has a mean quality rating that is significantly different from the others . Figure 7 summarises the analysis visually.

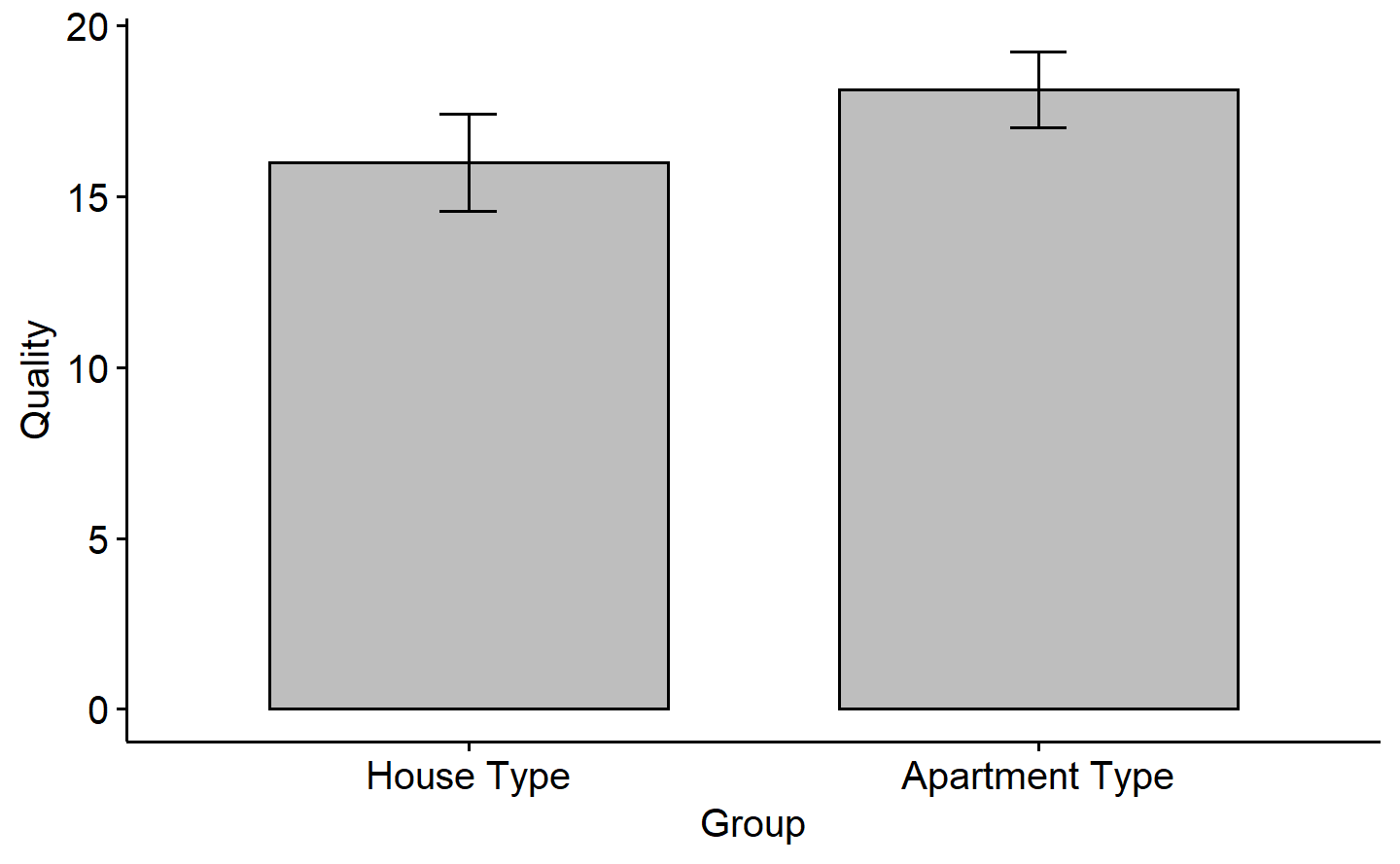
We then proceed to test our own hypotheses through planned comparisons. To compensate for multiple comparisons, we use the Bonferroni adjusted -values. The result of the first comparison suggests that there is no significant difference in the quality of Home-type sitcoms compared to Workplace-type sitcoms , . This is seen in figure 8. The second comparison reveals that Apartment-type sitcoms were rated as having significantly better quality than House-type sitcoms , . This is illustrated in figure 9. The third comparison reveals that the quality of Seinfeld is rated significantly lower than that of Friends , . This can be seen in figure 10. Finally, the fourth comparison suggests that the quality of Brooklyn Nine-Nine is rated significantly higher than that of The Office , . Figure 11 summarizes this visually.



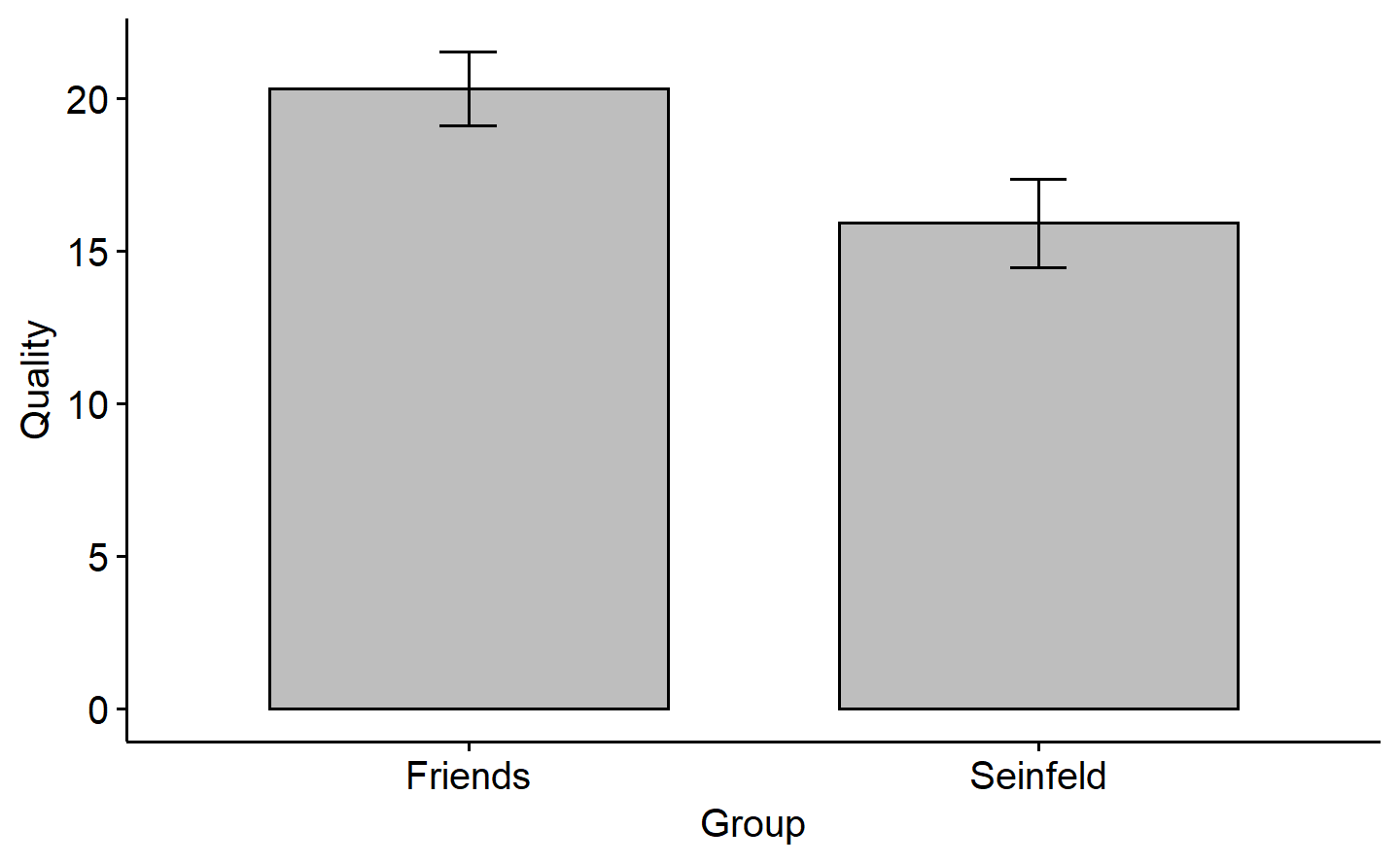
*Figure* *7.*  bar-plot comparing mean sitcom quality across all sitcoms



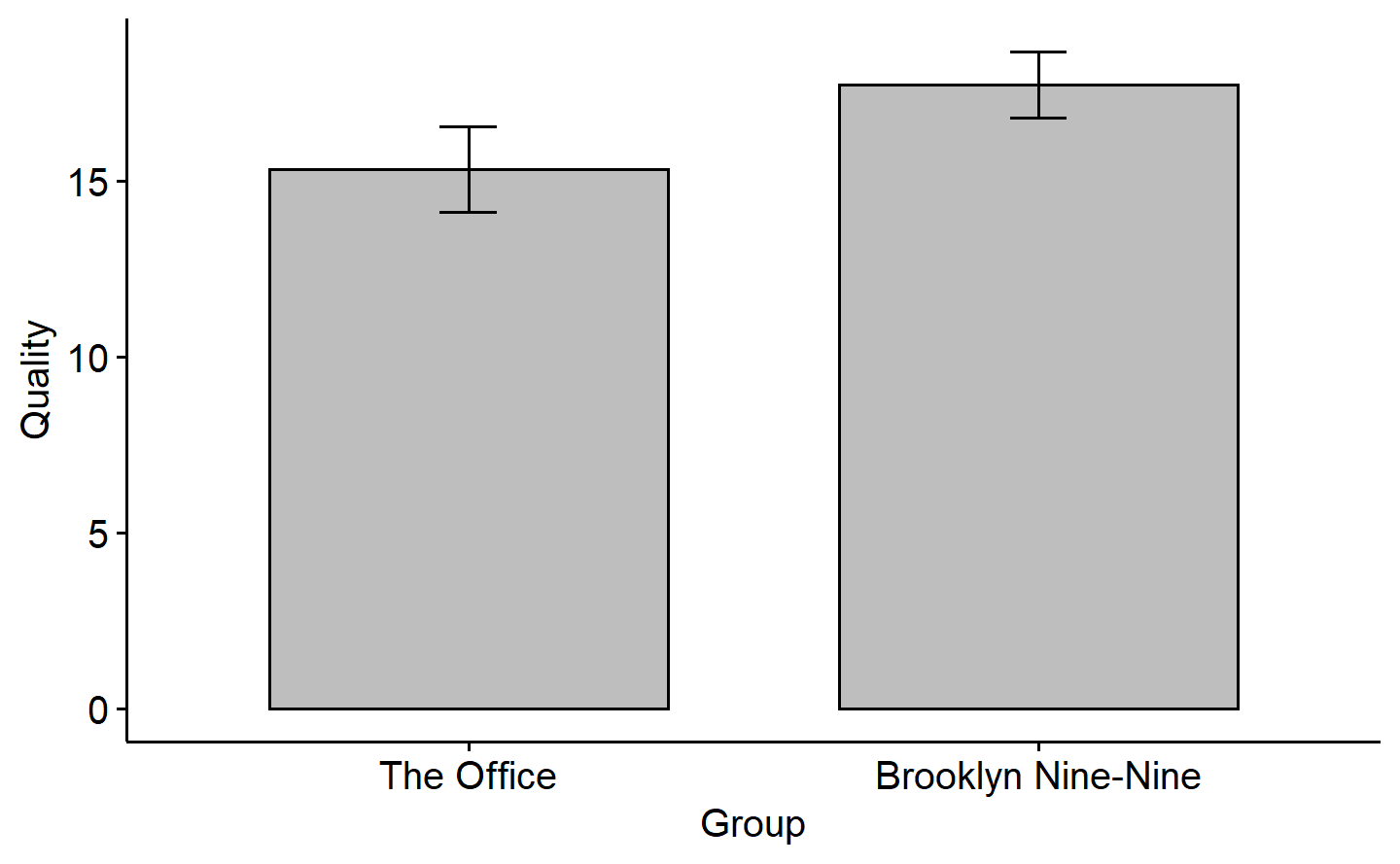
*Figure* *8.*  Bar-plot comparing mean sitcom quality between Home type and Workplace type sitcoms



*Figure* *9.*  Bar-plot comparing mean sitcom quality between House type and Apartment type sitcoms



*Figure* *10.*  Bar-plot comparing mean sitcom quality between Seinfeld and Friends



*Figure* *11.*  Bar-plot comparing mean sitcom quality between The Office and Brooklyn Nine-Nine

## Question 3d

*A graduate student who participated in your study liked your sitcom research idea and wants to use your same study design but doesn’t like the planned comparisons you chose and wants you to provide suggestions on what they could do (apparently they don’t want to do any work). Without repeating any of the three planned comparisons you originally did for your hypotheses, what are TWO complete/full orthogonal sets (distinct from each other as well), in both words and contrast codes, that this other graduate student could do? (2 marks)*

*Note 1. The actual comparisons themselves need to be different, not just the codes. For example, (1,-1,0) is not different from (2,-2,0) or (-1,1,0).*

*Note 2. You don’t need any kind of home/apartment theory to justify your comparisons here. It is just about orthogonality!*

Answers here.

First set:

contrast 1 = [-1,-1,-1,-1,4] (Brooklyn Nine-Nine is better than all other sitcoms)

contrast 2 = [-1,-1,-1,3,0] (The Office is better than all Home type sitcoms)

contrast 3 = [-1,-1,2,0,0] (Seinfeld is better than Modern Family and Friends)

contrast 4 = [-1,1,0,0,0] (Friends is better than Modern Family)

Second set:

contrast 1 = [-1,-1,4,-1,-1] (Seinfeld is better than all other sitcoms)

contrast 2 = [-1,-1,0,3,-1] (The Office is better than Modern Family, Friends, and Brooklyn Nine-Nine)

contrast 3 = [-1,2,0,0,-1] (Friends is better than Modern Family and Brooklyn Nine-Nine)

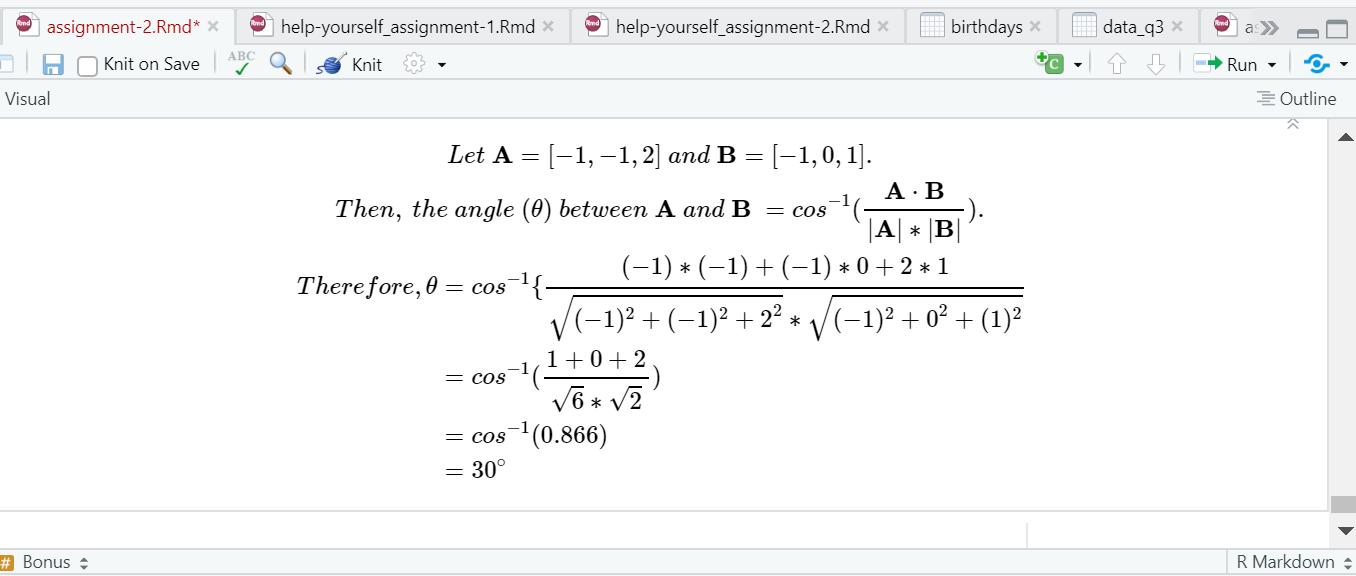
contrast 4 = [-1,0,0,0,1] (Brooklyn Nine-Nine is better than Modern Family)

# Bonus

*You have two planned comparisons’ vectors: [-1,-1,2] and [-1,0,1]. What is the angle between these two vectors (in degrees) and what does this mean for their variance? (.5 BONUS mark)*

$$ Let\ \mathbf{A} = [-1,-1,2]\ and\ \mathbf{B}=[-1,0,1].\\ Then,\ the \ angle \ (\theta) \ between \ \mathbf{A} \ and \ \mathbf{B} \ = cos^{-1}(\frac{\mathbf{A}\cdot \mathbf{B}}{|\mathbf{A}|\*|\mathbf{B}|}).\\ $$

**[Note: The output of this mathematical expression appears perfectly in the R window (attached screenshot), but does not show up on knitting the markdown document. Within the same document, however, another implementation in Q3a does show the output, although they are not aligned here, but were in the window]**



The fact that the angle between the the two contrast vectors is not radians means that they are not orthogonal. This means that the variance explained by the two contrasts is overlapping, and not independent.