**Cleaned Data Wrangling/Tidying Code for R**

1. Begin work in R, set working directory to desktop (where data file currently lives), install and use the packages “dplyr” and “tidyr” to use throughout the exercise, and read in data file to use (titled here “fpet”, or Ficus petiolaris nematode and fig wasp data).

setwd("C:/Users/justi/Desktop")

install.packages("dplyr")

library(dplyr)

install.packages("tidyr")

library(tidyr)

fpet <- read.csv("C:/Users/justi/Desktop/fpet.csv")

View(fpet)

2. Use the tidyr function “unite” to combine data columns for year, site, and tree to have a specific identifier for each fig. Save the changes that are made by titling new file.

fpet2 <- unite(fpet, "FigIdentifier", c("Year", "Site", "Tree"))

3. Count data was originally collected for males and female wasps of each species. Total counts per species can be summed with “mutate” function in dplyr.

fpet3 <- mutate(fpet2, PegoTot = PegoF+PegoM, IdarnesTot = LO1+SO1+SO2+Idmale, Het1Tot = Het1+Het1MW+Het1MU, Het2Tot = Het2+Het2MW+Het2MU, HetTot = Het1+Het1MW+Het1MU+Het2+Het2MW+Het2MU, AepoTot = AepoF+AepoM, PhysoTot = PhysoF+PhysoM)

4. A variable number of pollinating foundresses lay their eggs and pollinate before dying within the fig. Sometimes we sampled figs that were not visited by a foundress, and as such no pollinating offspring were produced. To exclude these 0 foundress figs from the dataset I used the dplyr function “filter”.

fpet4 <- filter(fpet3, Found > 0)

Further, I used the same function to eliminate any rows in which there were no pollinating offspring for other reasons. This eliminated over 700 rows.

fpet5 <- filter(fpet4, PegoTot > 0)

5. To ensure that no duplicate rows were used in this dataset I used the dplyr function “distinct”. There weren’t any duplicate rows, thankfully.

Fpet6 <- distinct(fpet5)

6. To find the mean foundress count for *F. petiolaris* I used the summarize function. This was 1.437.

summarize(fpet6, avg = mean((Found))

7. To order the number of rows for pollinator offspring from lowest to highest I used the arrange function.

Fpet7 <- arrange(fpet6, PegoTot)

8. To visualize (very roughly) the difference between the pollinator offspring production in nematode infested and uninfested figs I used a bar graph function in ggplot2.

ggplot(data=fpet8, aes(x=Nema, y=PegoTot))+geom\_bar(stat="identity")

The plot (rough as it is) is as follows:

