CONS 452, R-Lab 2
Marina Galvao – [redacted]
January 28th, 2022

1) a) Assuming that the NA values are real, the mean for nitrogen costs is 97.156, and the standard deviation is 55.69243.

b) After changing these values to 0.00 as opposed to removing them, there are five rows which possess a 0.00 under nitrogen costs. The new mean is 83.75517, and the new standard deviation is 61.81447. These differences are highly significant, the mean is lowered given that 0's are present, and the standard deviation presented becomes higher.

I believe these values are NA instead of a 0.00. To begin, all the values that possess NA are soybean, so it is probable that there may be something about the nature of soybean which inhibited nitrogen cost data collection under these specific soybean scenarios. Furthermore, those same rows also possess NA values for the "Drying" and "Price.Mg" columns, and the likelihood of all three of these being 0.00 as opposed to NA is very small. Assuming 0.00 significantly altered the mean (from 97.156 to 83.75517) and is a statement which is likely incorrect. Lastly, other values of soy nitrogen costs were nowhere near 0.00, which further proves that the true value is likely NA.

2) Costs$Price.Mg[is.na(Costs$Price.Mg)] = 185.2

3) Mean Total Cost variable = 959.3983. Mean for Annual Revenue: = 2677.653. Standard Deviation for Annual Revenue: 1694.443

4) a) The mean Seed Cost for soybean is 138.5107.

b) The mean Seed Cost for corn is 249.8200.

The p-value is 3.187e-10, which is less than 0.05. Because the p is low, the null hypothesis is rejected and the concluding assumption is that the two groups are statistically significant. Furthermore, the t value is higher than 0 (t = 9.6237), providing further evidence that the cost of corn and the cost of soy are significantly different, given that a t-value of 0 indicates an acceptance of the null hypothesis.

5) As cost of labour increases, commodity price also increases (Figure 1). There is a positive linear correlation present here. This is evident from the statistical analysis values. Statistically speaking, the p-value was 2.069e-09, which concludes that the null hypothesis can be rejected, and the relationship is significant. The regression line is further evidence of the relationship present in the scatter plot, as numerous data points sit near the line. Furthermore, the multiple r-squared value is 0.7413, and adjusted r-squared (likely slightly more precise) value is 0.7317, providing further evidence of the significance of this relationship, given that this represents the goodness of fit for the dataset.

CONS 452, R-Lab 2
Marina Galvao – [redacted]
January 28th, 2022

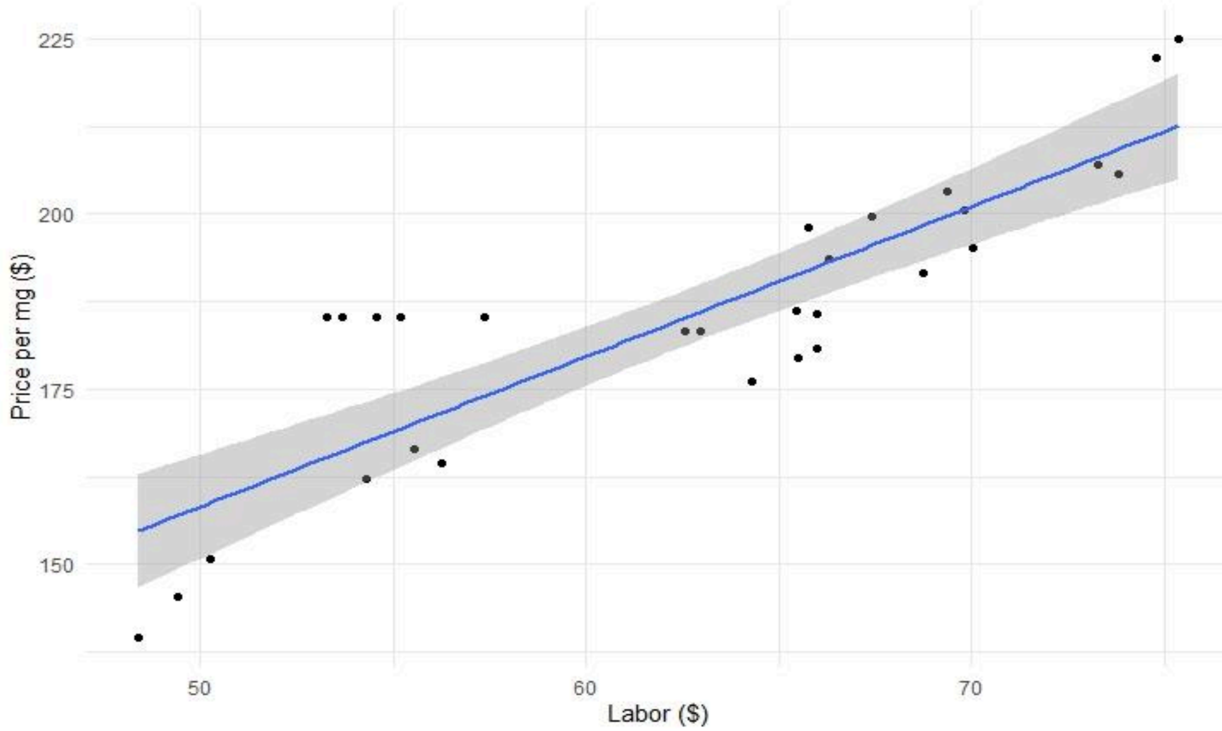

Figure 1, Positive linear relationship between Cost of Labor ($) and Price per mg ($). Cost of Labor is in the x-axis and Price per mg in the y-axis.

6) There is a linear relationship present for both crops, as can be observed in Figures 2 and 3. In the case of soy (Figure 2), this is a positive correlation where average yield (in mg) increases with herbicide costs. In the case of corn (Figure 3), average yield decreases as herbicide costs increase. This is likely because in the case of soy, the herbicide has a positive effect on the plants and soil, allowing it to grow better. Thus, the more spent on herbicides, the more successfully the soy grows. In the case of corn, the relationship is not as strong, and the data shows a slight decrease in average yield with increases in herbicide costs. In this case, it is likely that the herbicide does not impact the corn plantations. This is further proven by the p-values, as the relationship for corn yield and herbicide costs is not statistically significant, with a p-value greater than 0.05 (a value of 0.533). Soybean on the other hand, had a p-value of 0.02011, showing the relationship is statistically significant.

CONS 452, R-Lab 2
Marina Galvao – [redacted]
January 28th, 2022

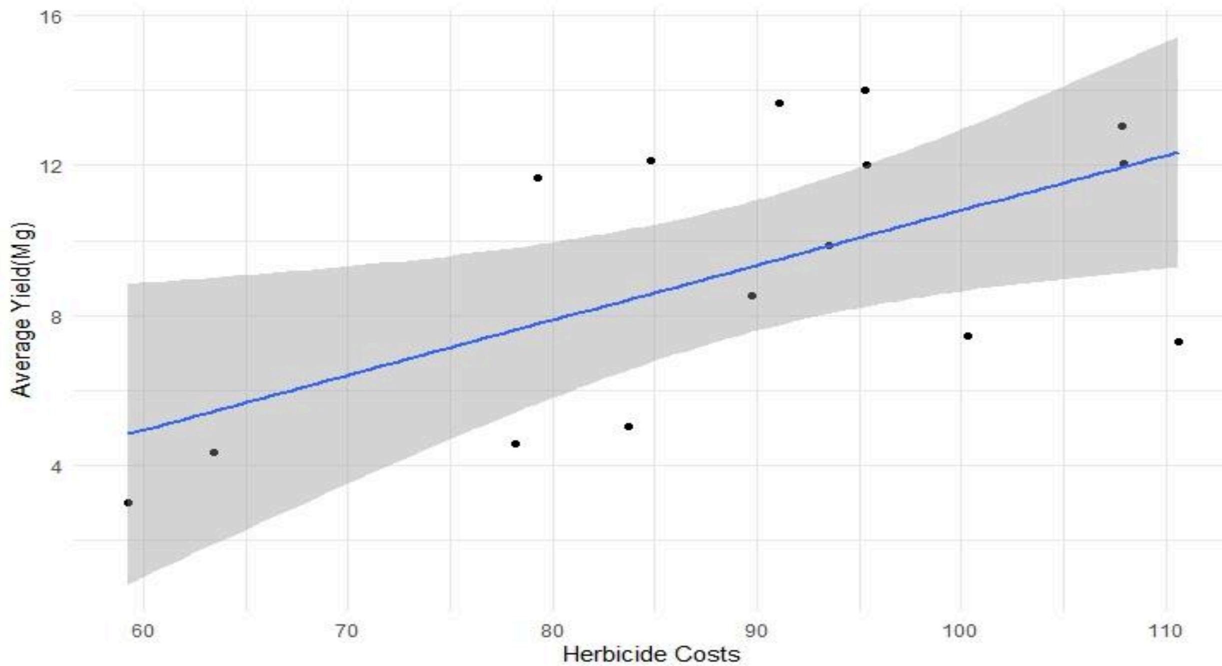

Figure 2, Herbicide Costs (in $) for soy in the x-axis, and Average Yield of soy in the y-axis (in mg). The relationship is a positive linear relationship, that which is statistically significant.

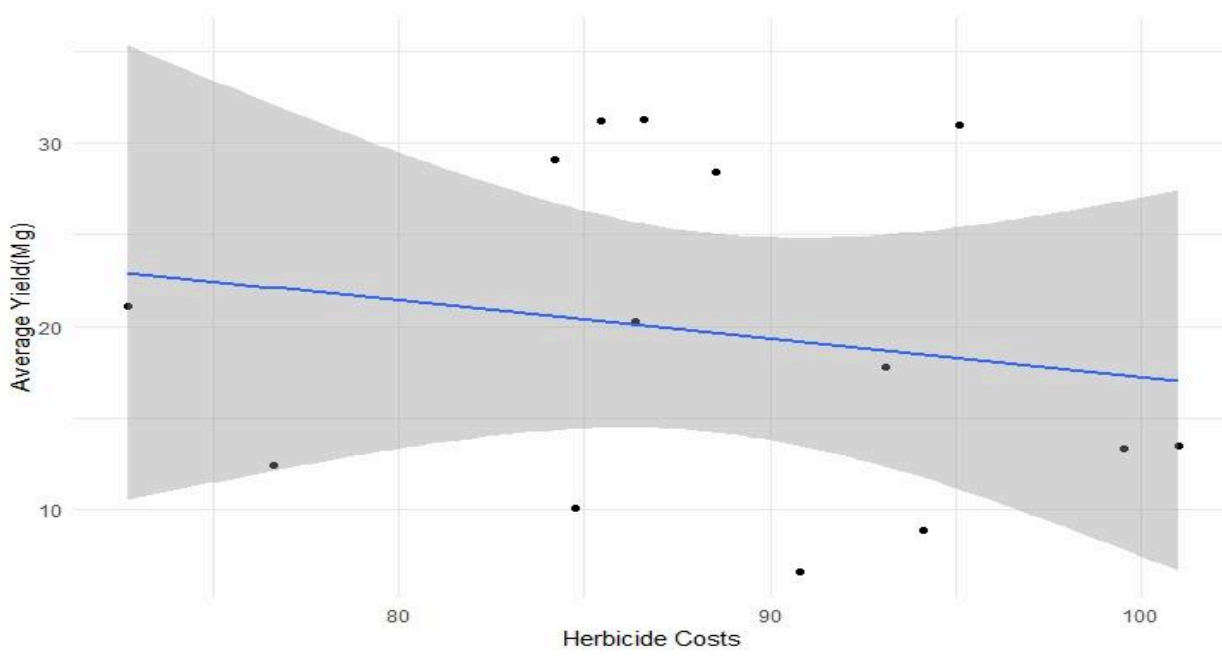

Figure 3, Herbicide Costs ($) for corn in the x-axis, and Average Yield of corn in the y-axis (in mg). This linear relationship is not statistically significant.

CONS 452, R-Lab 2
Marina Galvao – [redacted]
January 28th, 2022

7) The three treatments present were "Row C: The conventional full row cropping system", "0.10PS: Row crops with 10% prairie strips added in", and "0.20PS" Row crops with 20% prairie strip coverage". My hypothesis is that ecosystem services will increase as percentage of prairie strips increases. In other words, the site with the overall highest levels of ecosystems services will be 0.20PS, followed by 0.10PS, and Row C. I hypothesize this because prairie strips may successfully enhance biodiversity, which often possesses a positive correlation with other ecosystem services.

In the ANOVA tests performed, the null hypothesis describes no significant difference amongst group means. In tables 2-6 (shown in appendix), all p-values recorded are less than 0.05, meaning the null hypothesis can be rejected and a statistically significant difference is present amongst all group means.

Comparing the three different treatments with varying ecosystem services (Figures 4 to 8), a comparison to pollinator abundance (Figure 4) showed the same patterns as bird species richness (Figure 7) – with the highest level of the ecosystem service being 0.20PS, followed by 0.10PS and RowC. This can be explained by the often-positive correlation observed between pollinator species and bird species. Positive biodiversity effects often trickle down along an entire ecosystem. Moreover, some pollinators such as hummingbirds are considered bird species as well, which may be an underlying factor behind the relationship present. Lastly, the order of ecosystem services is in line with the hypothesis proposed (0.20PS, 0.10PS, and RowC).

Two other similar figures were surface runoff (Figure 6) and crop yield richness (Figure _8), with the highest ecosystem service level being crop type RowC, followed by 0.10PS, and 0.20PS. This result does not align with the hypothesis I presented. The reason this does not align might be that my hypothesis only took into consideration positive impacts of ecosystem services, however, large quantities of surface runoff are considered an ecosystem service, albeit this action is unlikely to have a positive impact to this ecosystem. The same logic can be used for crop yield, as too high of a crop yield can negatively impact a system.

One explanation for this relationship present may be that the more intensive the crop plantation/the higher the crop yield, the more detrimental results such as surface runoff will be present. It appears that it is likely that crop yield and surface runoff possess an inverse relationship to bird species richness and pollinator abundance, with the first two benefiting from intense agricultural practices, and while the last two are negatively impacted

Insect taxa richness by site (Figure 5) shows patterns different to the rest, with 0.10PS presenting the highest level of ecosystem service, just slightly under is 0.20PS, and followed by RowC. In this case, given how similar 0.10PS and 0.20PS show on the graph, it may be possible that insect taxa richness possesses a threshold under which it no longer significantly grows without further changes. It is worth noting that RowC still possesses the lowest level of this service, as stated in part of my hypothesis.

CONS 452, R-Lab 2
Marina Galvao – [redacted]
January 28th, 2022

All ecosystem services in this study are significantly different amongst crop types. In addition, a Two-Sample t-test was performed to compare Pollinator Abundance between the site 0.10PS and RowC. In this case, the alternative hypothesis is that "the true difference in means between group 0.10PS and group RowC is not equal to 0." In comparing pollinator abundance by Crop Type, the p-value of 0.0004398 shows the alternative hypothesis can be accepted and the null hypothesis rejected, resulting in a significant difference. In this case, the mean in group 0.10PS is 20.56704, and in RowC, 11.85714. The difference in these means (with 0.10PS being significantly higher than RowC) and their significant differences further corroborates my hypothesis that ecosystem services increase as percentage of prairie strips increases. These results support the use of prairie strips in intensive agriculture systems, as positive ecosystem services were proven to be enhanced.

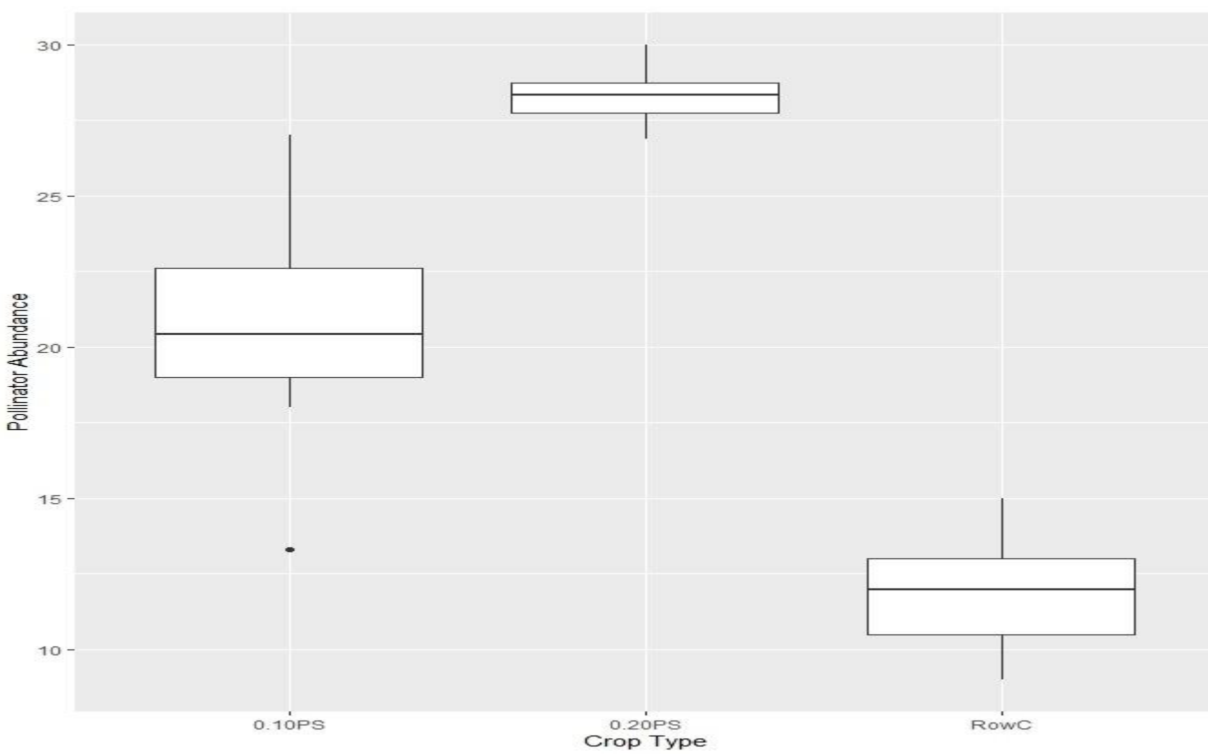

Figure 4, Level of Pollinator Abundance (y-axis) in relation to Crop Type (x-axis, options include RowC, 0.20PS, and 0.10PS). The interquartile range is highest in 0.10PS, followed by RowC and 0.20Ps.0.20PS possesses the highest Pollinator Abundance, followed by 0.10PS, and RowC. The only outlier present can be seen is 0.10PS, that which also possesses the longest highest maximum range.

CONS 452, R-Lab 2
Marina Galvao – [redacted]
January 28th, 2022

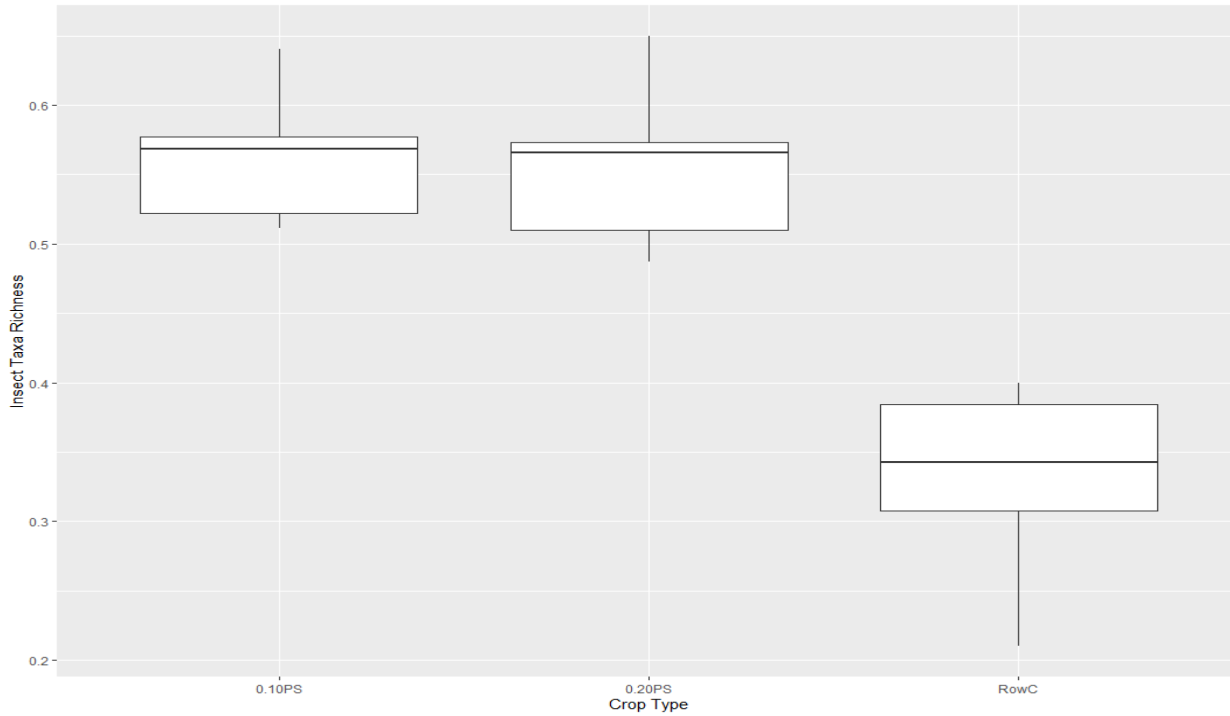

Figure 5, Level of Insect Taxa Richness (y-axis) in relation to Crop Type (x-axis, options include RowC, 0.20PS, and 0.10PS). The Insect Taxa Richness is highest in 0.10PS, followed closely behind by 0.20PS, which is followed by RowC. RowC possesses the highest interquartile range and the lowest minimum value. There are no outliers present.

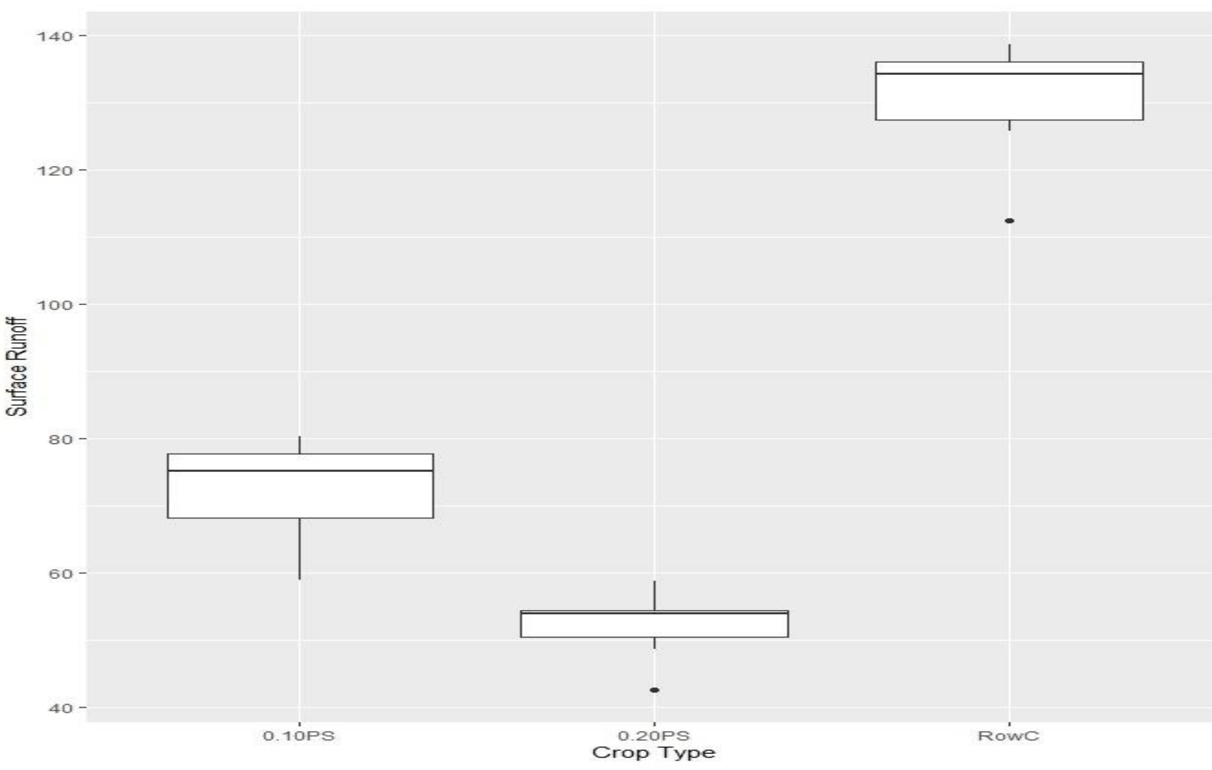

Figure 6, Level of Surface Runoff (y-axis) in relation to Crop Type (x-axis, options include RowC, 0.20PS, and 0.10PS). RowC possesses the highest level of Surface Runoff, followed by 0.10PS and 0.20PS. The interquartile range is highest in 0.10PS, followed by RowC and 0.20Ps. There is an outlier present in 0.20PS and in RowC.

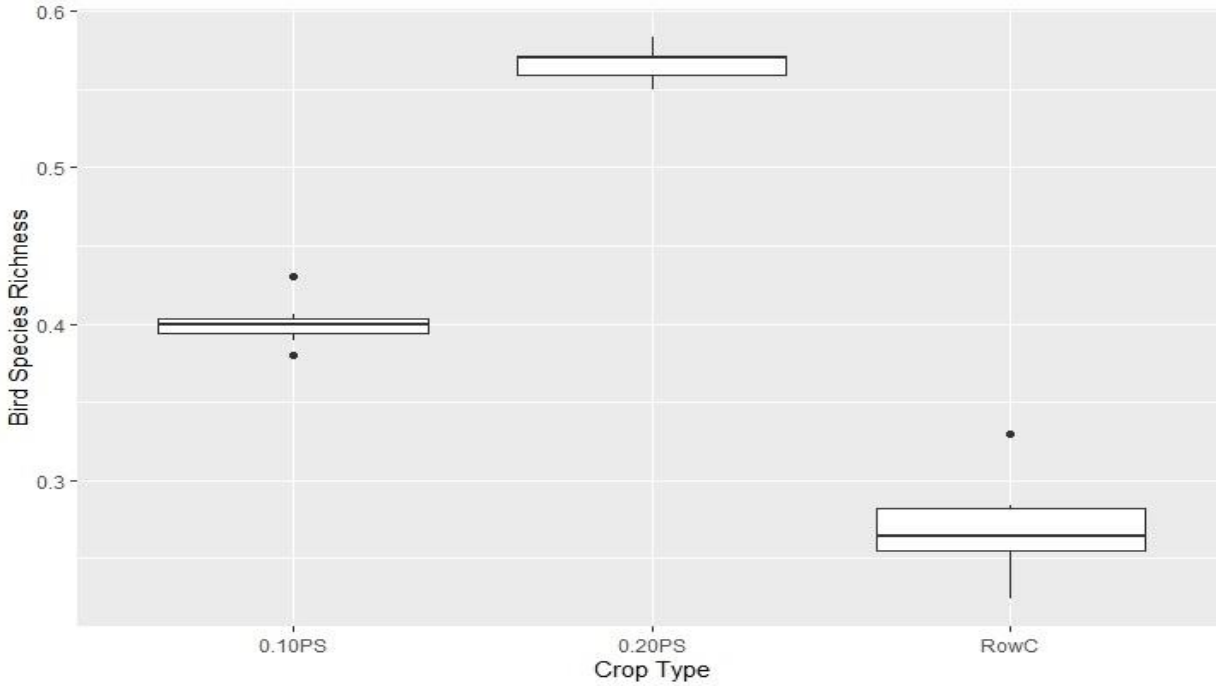

Figure 7, Level of Bird Species Richness (y-axis) in relation to Crop Type (x-axis, options include RowC, 0.20PS, and 0.10PS). 0.20PS possesses the highest level of Bird Species Richness, followed by 0.10PS and RowC. The interquartile range is highest in RowC, which has the greatest minimum value, and that which possesses an outlier. There are also two outliers present in 0.10PS.

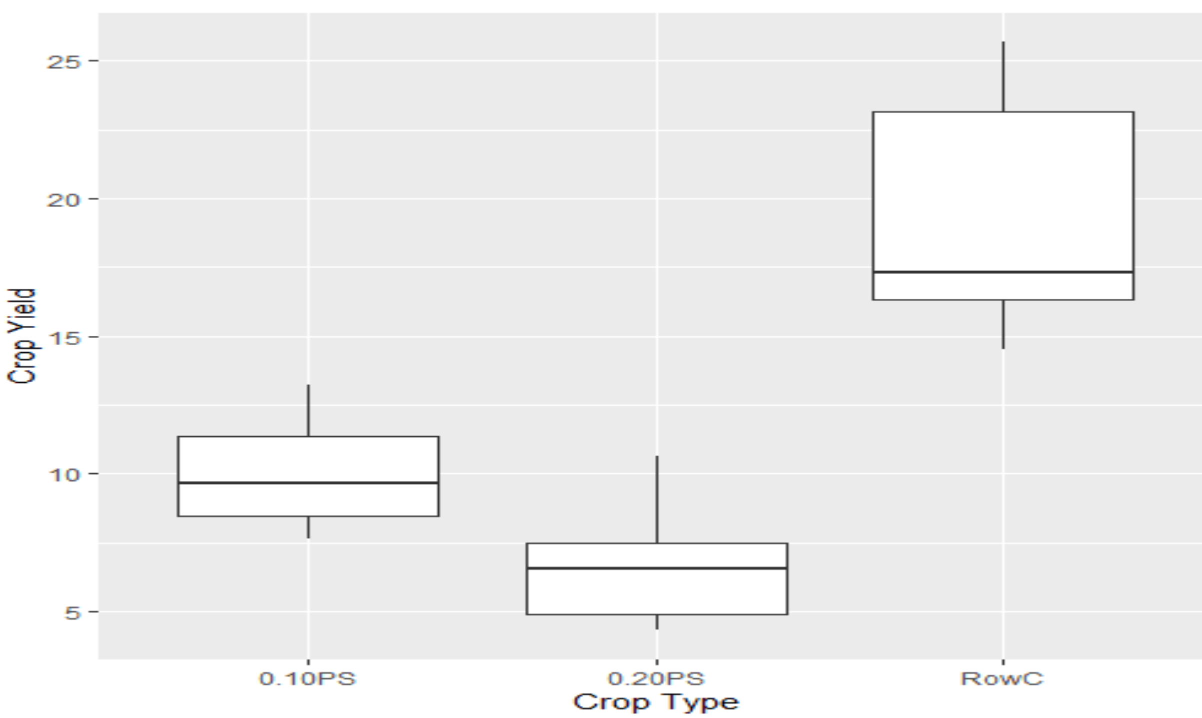

Figure 8, Level of Crop Yield (y-axis) in relation to Crop Type (x-axis, options include RowC, 0.20PS, and 0.10PS). RowC possesses the highest level of Crop Yield, followed by 0.10PS and 0.20Ps. The interquartile range is highest in RowC. There are no outliers present.

CONS 452, R-Lab 2
Marina Galvao – [redacted]
January 28th, 2022

8) The ecosystem services I chose were Surface Runoff and Crop Yield. All information can be found under the t-tests in the r-script file, as well as the appendix. Comparisons of 0.10PS, 0.20PS and RowC were performed via t-tests. The results are shown below.

Table 1, Summary of six t-tests performed to compare differences in Surface Runoff and Crop Yield between crop types 0.10PS, 0.20PS, and RowC.

	p-value of Surface Runoff	p-value of Crop Yield
0.10PS and RowC	3.141e-08	0.0002782
0.10 PS and 0.20PS	2.599e-05	0.01033
0.20PS and RowC	2.33e-12	3.016e-06

Table 1 shows evidence that all of the t-tests resulted in statistically significant p-values, given that all values are below 0.05. Overall, the highly important statistically significant presence here coupled with the figures imply synergy between pollinator abundance (Figure 4) and bird species richness (Figure 7), and synergies between surface runoff (Figure 6) and crop yield (Figure 8). Insect taxa does not show the same pattern as the rest, and the possible reasons for this were discussed in the answer to question 7. In relation to trade-offs, the data implies that there are trade-offs present in levels of pollinator abundance and bird species richness in relation to levels of surface runoff and crop yield.

Agricultural practices are often coupled with ecosystem deterioration and biodiversity decreases, however, this study showed the possibility of increasing biodiversity levels while maintaining plantations. Overall, the implementation of crop types 0.20PS and 0.10PS is extremely wise, crop yields are not significantly altered and increases in positive ecosystem services are evident.

The integration of these strips amongst the plantations was shown to increase biodiversity (Figures 4 and 7), as well as increase insect taxa richness (Figure 5), which showed higher levels in ecosystems 0.20PS and 0.10PS in comparison to RowC. Lastly, the increases in surface runoff shown (Figure 6) imply a greater water retention time in crop types 0.20PS and 0.10PS, that which positively impacts the plantation and has positive correlations with other ecosystem services (soil biodiversity, decreased soil erosion rates, amongst others). It is also worth noting that it is likely that high surface runoff values seen in RowC (Figure 6) negatively impact insects inhabiting the soil or acquiring resources from the soil. The trade-offs between the positive effects (pollinator abundance, bird species richness and insect taxa richness) and the negative effects (surface runoff) are proof of the beneficial results of the strips. This conclusion can be further corroborated by the fact that the synergies present here will enhance positive and negative effects.

CONS 452, R-Lab 2
Marina Galvao – [redacted]
January 28th, 2022

Appendix

Table 2, Analysis of Variance Table for Pollinator Abundance between the three sites

Response: Pollinator Abundance

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
CropType	2	1072.26	536.13	72.287	7.026e-10	***
Residuals	20	148.33	7.42			

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 3, Analysis of Variance Table for Pollinator Abundance between the three sites

Response: Insect Taxa Richness

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
CropType	2	0.246591	0.123295	38.495	1.712e-07	***
Residuals	20	0.065767	0.003288			

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 4, Analysis of Variance Table for Pollinator Abundance between the three sites

Response: Surface Runoff

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
CropType	2	24865.6	12432.8	230.48	1.546e-14	***
Residuals	20	1078.8	53.9			

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 5, Analysis of Variance Table for Bird Species Richness between the three sites

Response: Bird Species Richness

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
CropType	2	0.35281	0.176405	393.55	< 2.2e-16	***
Residuals	20	0.00896	0.000448			

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 6, Analysis of Variance Table for Crop Yield between the three sites

Response: Crop Yield

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
CropType	2	668.89	334.45	35.459	2.653e-07	***
Residuals	20	188.64	9.43			

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

CONS 452, R-Lab 2
Marina Galvao – [redacted]
January 28th, 2022

T-Tests:

```
> t.test(Surface_Runoff ~ CropType, data = T_test1, var.equal = TRUE)

	Two Sample t-test

data:  Surface_Runoff by CropType
t = -12.473, df = 12, p-value = 3.141e-08
alternative hypothesis: true difference in means between group 0.10PS and group RowC is not equal to 0
95 percent confidence interval:
 -68.07949 -47.83134
sample estimates:
mean in group 0.10PS   mean in group RowC
        72.36332           130.31874

> t.test(Crop_Yield ~ CropType, data = T_test1, var.equal = TRUE)

	Two Sample t-test

data:  Crop_Yield by CropType
t = -5.0635, df = 12, p-value = 0.0002782
alternative hypothesis: true difference in means between group 0.10PS and group RowC is not equal to 0
95 percent confidence interval:
 -13.559223  -5.400777
sample estimates:
mean in group 0.10PS   mean in group RowC
        10.02143            19.50143

> T_test2 = subset(PrairieStrips_ES, PrairieStrips_ES$CropType == "0.10PS" | PrairieStrips_ES$CropType == "0.20PS")
> t.test(Surface_Runoff ~ CropType, data = T_test2, var.equal = TRUE)

	Two Sample t-test

data:  Surface_Runoff by CropType
t = 6.1315, df = 14, p-value = 2.599e-05
alternative hypothesis: true difference in means between group 0.10PS and group 0.20PS is not equal to 0
95 percent confidence interval:
 12.96575 26.91637
sample estimates:
mean in group 0.10PS mean in group 0.20PS
        72.36332             52.42226
```

CONS 452, R-Lab 2
Marina Galvao – [redacted]
January 28th, 2022

```
> t.test(Crop_Yield ~ CropType, data = T_test2, var.equal = TRUE)

	Two Sample t-test

data:  Crop_Yield by CropType
t = 2.9602, df = 14, p-value = 0.01033
alternative hypothesis: true difference in means between group 0.10PS and group 0.20PS is not equal to 0
95 percent confidence interval:
 0.9097491 5.6953303
sample estimates:
mean in group 0.10PS mean in group 0.20PS
         10.021429             6.718889

> T_test3 = subset(PrairieStrips_ES, PrairieStrips_ES$CropType == "0.20PS" | PrairieStrips_ES$CropType == "RowC")
> View(T_test3)
> t.test(Surface_Runoff ~ CropType, data = T_test3, var.equal = TRUE)

	Two Sample t-test

data:  Surface_Runoff by CropType
t = -22.38, df = 14, p-value = 2.33e-12
alternative hypothesis: true difference in means between group 0.20PS and group RowC is not equal to 0
95 percent confidence interval:
 -85.36164 -70.43131
sample estimates:
mean in group 0.20PS   mean in group RowC
          52.42226           130.31874

> t.test(Crop_Yield ~ CropType, data = T_test3, var.equal = TRUE)

	Two Sample t-test

data:  Crop_Yield by CropType
t = -7.4693, df = 14, p-value = 3.016e-06
alternative hypothesis: true difference in means between group 0.20PS and group RowC is not equal to 0
95 percent confidence interval:
 -16.453013  -9.112066
sample estimates:
mean in group 0.20PS   mean in group RowC
          6.718889           19.501429
```