

# Biodiversity Sampling Final Group Submission

Lucy, Hannah, Gena, Paige

01/11/2021

Loading in data:

Organizing data & making new columns:

```
merged_data_1 = merge(biodiversity, cover)
```

```
merged_data_1 <- merged_data_1 %>%  
  select(-c(biotic_hab, per_mussel, per_barnacle, per_upright, per_crust, num_snail, num_crab))
```

*#number of individuals per species in the 10 squares*

```
quad_shannon <- biodiversity %>%  
  drop_na(n) %>%  
  group_by(vertical_transect, quadrat, species) %>%  
  summarize(n_per_species = sum(n))
```

## 'summarise()' has grouped output by 'vertical\_transect', 'quadrat'. You can override using the '.groups' argument

*#of occurrences per quadrat*

```
quad_occ_count <- biodiversity %>%  
  drop_na(species, n) %>%  
  group_by(vertical_transect, quadrat, species) %>%  
  summarize(occ_count_per_quadrat = n())
```

## 'summarise()' has grouped output by 'vertical\_transect', 'quadrat'. You can override using the '.groups' argument

*#of species per quadrat*

```
quad_species_count <- quad_occ_count %>%  
  group_by(vertical_transect, quadrat) %>%  
  summarize(species_per_quad = n())
```

## 'summarise()' has grouped output by 'vertical\_transect'. You can override using the '.groups' argument

*#total number of individuals of all species in quadrat found in the 10 squares*

```
quad_total <- biodiversity %>%  
  drop_na(species, n) %>%  
  group_by(vertical_transect, quadrat) %>%  
  summarize(n_total_per_quadrat = sum(n))
```

## 'summarise()' has grouped output by 'vertical\_transect'. You can override using the '.groups' argument

```
quad_shannon_2 <- merge(quad_shannon, quad_total, all = TRUE, by = c('vertical_transect', 'quadrat'))
```

```
quad_shannon_3 <- quad_shannon_2 %>%  
  drop_na(species) %>%  
  mutate(p_i = n_per_species/n_total_per_quadrat) %>%  
  mutate(Pi_lnPi = p_i*log(p_i))
```

*#calculating shannon wiener for each quadrat and adding to merged data*

```
shannon_index <- quad_shannon_3 %>%  
  group_by(vertical_transect, quadrat) %>%  
  summarize(H = -sum(Pi_lnPi))
```

## 'summarise()' has grouped output by 'vertical\_transect'. You can override using the '.groups' argument

```
merged_data_2 <- merge(quad_shannon, cover)
merged_data_3 <- merge(merged_data_1, shannon_index)
```

Then we organize the final dataset by merging the above, made code for plots and did statistical analysis using linear models

```
##
## Call:
## lm(formula = per_cobble ~ species_per_quad, data = figure_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -48.029 -23.614  -3.029   27.410   49.141
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    22.2748     3.7617   5.921 7.52e-09 ***
## species_per_quad  8.5847     0.9938   8.638 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 28.58 on 356 degrees of freedom
## Multiple R-squared:  0.1733, Adjusted R-squared:  0.171
## F-statistic: 74.61 on 1 and 356 DF, p-value: < 2.2e-16
```

For species richness corresponding to %cobble cover:  $p = 2e-16$ , significant

```
##
## Call:
## lm(formula = per_cobble ~ H, data = figure_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -51.801 -20.985  -4.419   23.954   47.256
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    56.046     4.101   13.67 <2e-16 ***
## H             -4.721     4.413   -1.07   0.285
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 31.38 on 356 degrees of freedom
## Multiple R-squared:  0.003205, Adjusted R-squared:  0.0004052
## F-statistic: 1.145 on 1 and 356 DF, p-value: 0.2854
```

#For Shannon Wiener Index corresponding to %cobble cover:  $p = 0.285$ , not significant

```
##
## Call:
```

```
## lm(formula = per_pebble ~ H, data = figure_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -21.283 -16.412   2.113  12.408  30.752
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    4.255      2.122   2.006  0.0457 *
## H              17.384      2.283   7.614 2.42e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16.24 on 356 degrees of freedom
## Multiple R-squared:  0.14, Adjusted R-squared:  0.1376
## F-statistic: 57.97 on 1 and 356 DF, p-value: 2.42e-13
```

For Shannon Wiener Index corresponding to %pebble cover: p-value = 2.42e-13, significant

## Biodiversity Sampling Assignment

*Question:* Does substrate type impact sessile invertebrate biodiversity in the intertidal zone?

Specifically, how does percent cobble cover impact species richness in each quadrat?, how does percent cobble cover impact the Shannon Wiener Index in each quadrat?, and how does percent pebble cover impact the shannon Wiener index in each quadrat?

*Data Collection:* We sampled our data at Scott's Bay, BC using transect tapes parallel to the shoreline and quadrats that were placed haphazardly along six vertical transects. Measurements were taken during a high low tide at two different elevations across exposed and sheltered sides of the bay, and species counts and percent biotic and abiotic cover was recorded for randomly selected squares in each quadrat.

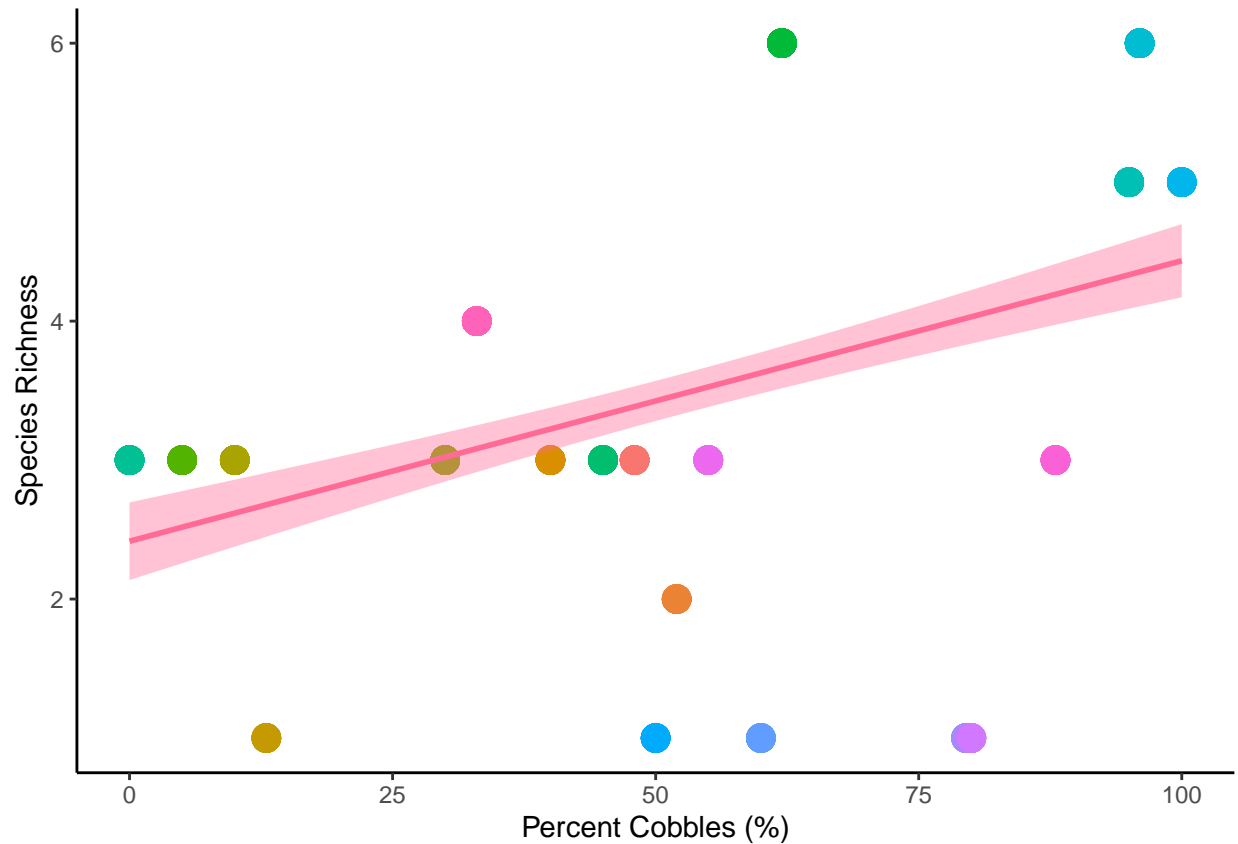
*Summary of Analysis:* We manipulated and organized our csv datasets in R studio to create a csv that included all data needed to create our figures. To answer our question we calculated species richness and Shannon Wiener Index for each quadrat, and then performed linear regression models to test the significance of our findings.

Parameters	p-value	Adjusted R-squared
Percent cobble vs Species Richness	< 2.2e-16	0.171
Percent cobble vs Shannon Weiner Index	0.2854	0.0004052
Percent pebble vs Shannon Weiner Index	2.42e-13	0.1376

*Results & Conclusions:* We first plotted the species richness versus percent boulder cover in each quadrat, finding that boulder cover did significantly impact species richness in a positive correlation.

```
species_rich
```

```
## 'geom_smooth()' using formula 'y ~ x'
```

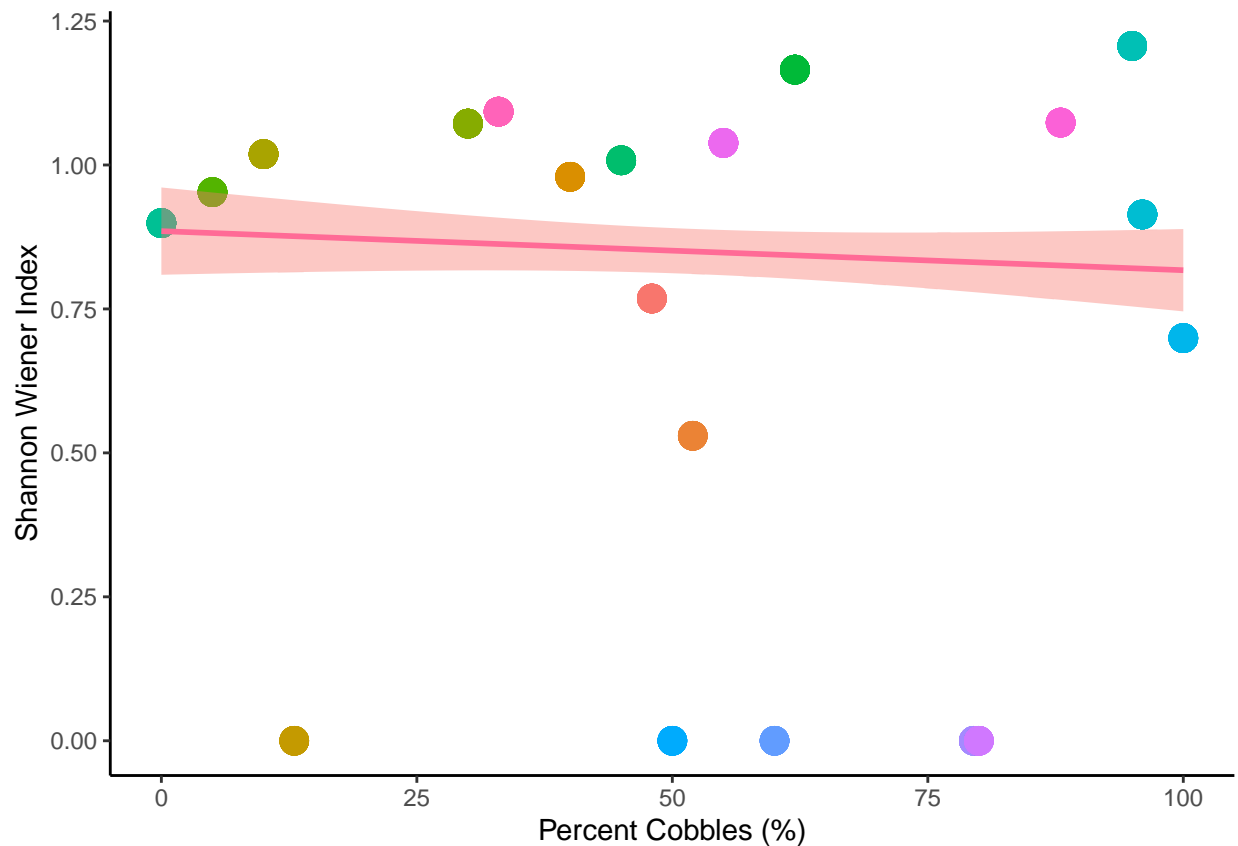


**Figure 1:** Species richness is the response variable plotted against percent cobbles, which serves as the explanatory variable. The colours represent different quadrats along our transects. The linear trendline was added and we found an adjusted  $R^2$  of 0.171 and a p-value of  $< 2.2e-16$ .

Then we plotted two figures for biodiversity (Shannon Wiener index) versus percent boulder and percent pebble cover, respectively, finding that cobble cover did not significantly impact biodiversity but pebble cover did have a significant positive correlation.

```
#for percent cobble  
shannon_plot_cobble
```

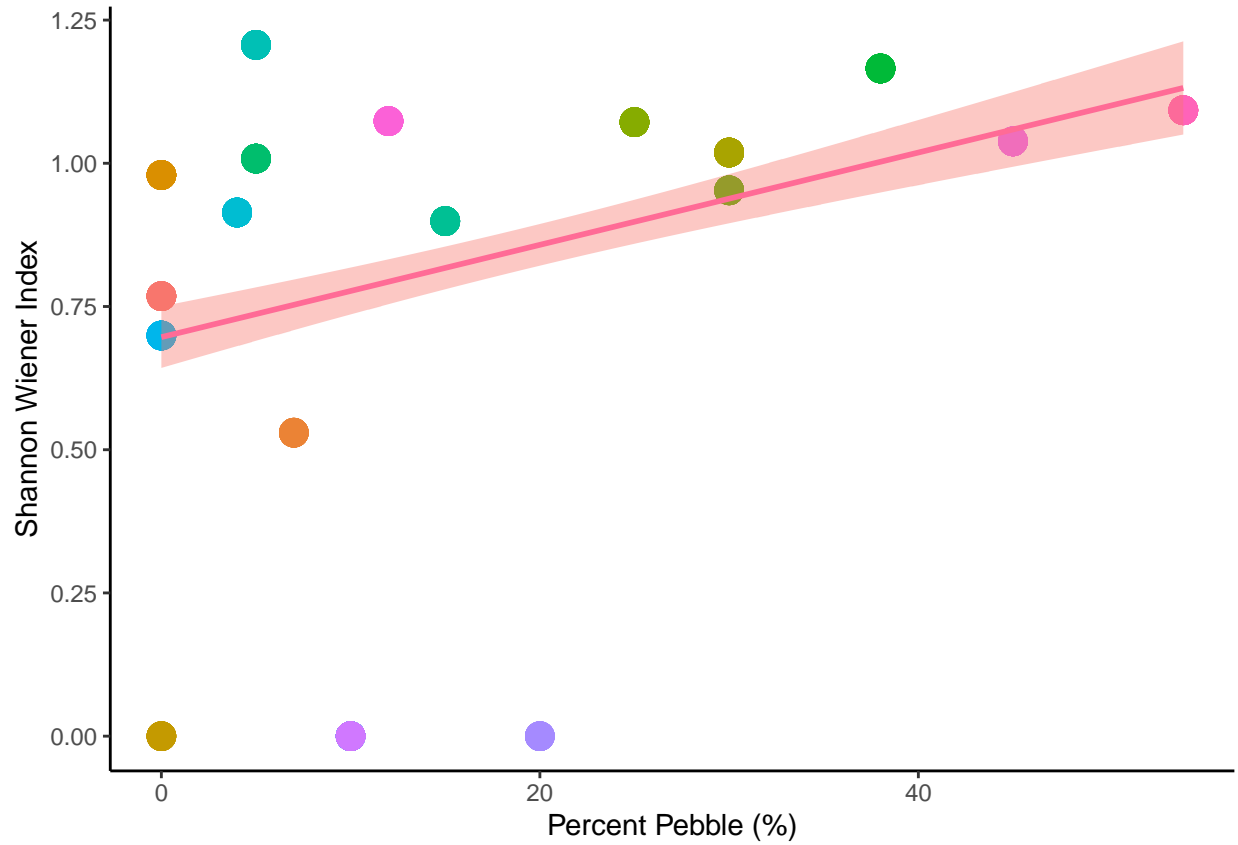
```
## 'geom_smooth()' using formula 'y ~ x'
```



**Figure 2:** The Shannon Wiener Index is the response variable plotted against percent cobbles, which serves as the explanatory variable. The colours represent different quadrats along our transects. The linear trendline was added and we found an adjusted  $R^2$  of 0.0004052 and a p-value of 0.2854.

```
#for percent pebble  
shannon_plot_pebble
```

```
## 'geom_smooth()' using formula 'y ~ x'
```



**Figure 3:** The Shannon Wiener Index is the response variable plotted against percent pebble, which serves as the explanatory variable. The colours represent different quadrats along our transects. The linear trendline was added and we found an adjusted  $R^2$  of 0.1376 and a p-value of 2.42e-13.