

# Monitoring Analysis Steps - Shrub vs Veg Cover

My goal is to compare the percentage of hits that are "shrub" to the percentage of hits that are "veg" on my plots.

This page will briefly cover the process of joins needed to easily read the data stored in the database. There is more detail in the Quadrat Species analysis page, the first steps are very similar.

In the SQL editor, use this code to create analysis table B1.

```
SELECT tlu_eJepson.genus, tlu_eJepson.species, tlu_eJepson.nativity
tblPointLineInterceptData.Dead,
tblPointLineInterceptData.Canopy,
tblPointLineInterceptData.TransectDistance,
tblPointLineInterceptData.EventID
INTO B1_2021
FROM tblPointLineInterceptData
LEFT JOIN tlu_eJepson
ON tblPointLineInterceptData.SpeciesID = tlu_eJepson.PlantSpeciesID;
```

This command joins data taken from the shrub transects to the Jepson manual records.

---

After you have created table B1, you will need to add data from the EventID table. This will include information about when and where the data was collected.

```
SELECT
tblEvents.transect,
tblEvents.LocationID,
tblEvents.EventGroup,
```

```

B1_2021.genus, B1_2021.species, B1_2021.nativity,
B1_2021.Dead,
B1_2021.Canopy,
B1_2021.TransectDistance

INTO
B2_2021

FROM
B1_2021

LEFT JOIN
tblEvents

ON
B1_2021.EventID = tblEvents.EventID;

```

This code performs a left join with tblEvents on the newly created B1 table.

---

The final join adds the text label for the EventGroup, which will show us the season that the data was taken.

```

SELECT
tblEventGroup.EventGroupName,
B2_2021.transect,
B2_2021.LocationID,
B2_2021.genus, B2_2021.species, B2_2021.nativity,
B2_2021.Dead,
B2_2021.Canopy,
B2_2021.TransectDistance

INTO B3_2021

FROM B2_2021

LEFT JOIN tblEventGroup

ON B2_2021.EventGroup = tblEventGroup.EventGroupID;

```

The final table B3 looks like this:

EventGroup	transect	LocationID	genus	species	nativity	Dead	Canopy	TransectDistance
Spring_2020	A	MGP_01_HWY	VEGETATION	VEGETATION	NOTPLANT	<input type="checkbox"/>	<input type="checkbox"/>	3
Spring_2020	A	MGP_01_HWY	VEGETATION	VEGETATION	NOTPLANT	<input type="checkbox"/>	<input type="checkbox"/>	3.25
Spring_2020	A	MGP_01_HWY	VEGETATION	VEGETATION	NOTPLANT	<input type="checkbox"/>	<input type="checkbox"/>	3.5
Spring_2020	A	MGP_01_HWY	Baccharis	pilularis	NATIVE	<input type="checkbox"/>	<input type="checkbox"/>	3.75
Spring_2020	A	MGP_01_HWY	Baccharis	pilularis	NATIVE	<input type="checkbox"/>	<input type="checkbox"/>	4
Spring_2020	A	MGP_01_HWY	Baccharis	pilularis	NATIVE	<input type="checkbox"/>	<input type="checkbox"/>	4.25
Spring_2020	A	MGP_01_HWY	VEGETATION	VEGETATION	NOTPLANT	<input type="checkbox"/>	<input type="checkbox"/>	4.5
Spring_2020	A	MGP_01_HWY	VEGETATION	VEGETATION	NOTPLANT	<input type="checkbox"/>	<input type="checkbox"/>	4.75

I exported this into an excel spreadsheet, then saved that exported spreadsheet as a CSV file.

```
import matplotlib.pyplot as plt
import pandas as pd
import seaborn as sns

df = pd.read_csv('B3_2021.csv')
df = df.replace(['UNKNOWN FRAGERIA'], ['VEGETATION'])
df = df.replace(['UNKNOWN LAMIACEAE'], ['VEGETATION'])
df = df.replace(['GALIUM'], 'VEGETATION')
df = df.replace(['UNKNOWN GALIUM'], ['VEGETATION'])
df = df.replace(['australis'], ['aurantiacus'])
df['Species'] = df['genus'] + ' ' + df['species']
df = df.replace(['LITTER LITTER'], ['Litter'])
df = df.replace(['VEGETATION VEGETATION'], ['Vegetation'])
dfalive = df[(df['Dead']==False)]
df01 = df[(df['LocationID']=='MGP_01_HWY')]
dfcount01 = df01.groupby(['EventGroupName', 'LocationID', 'Species'], as_index=False).size()
dfcount01 = dfcount01.rename(columns={'EventGroupName': 'Season'})
dfcount01['Percent Hits'] = round((dfcount01['size'] / 144) *100, 2)
dfcount01

sns.set(rc = {'figure.figsize':(17,3)})
sns.barplot(x = "Species", y = "Percent Hits", hue = "Season", data = dfcount01)
plt.legend(bbox_to_anchor=(.97, .95) )
plt.ylim(0, 100)
plt.title('MGP 01 HWY')
plt.show()
```

I relabeled the fragaria, lamiaceae, and galium hits as vegetation because they are not shrubs or vines. I combined the species and genus columns into one combined column. I relabeled the combined columns for the litter and vegetation hits to make them more legible. This code is discarding the live hits, a step for future analysis would be to find out how many locations had both live and dead hits of the same plant, which would indicate dieback. I found the percentage by dividing by the total number of hits at each plot (3 transects multiplied by 48 hits at each plot) and multiplying that total by 100. The code in the second paragraph creates the bar plot, specifying which data will be shown on which axis, the size of the plot, the location of the legend, and the title for the plot.