%Finding correlations in floral trait data

%red= red anthocyanin pigments, short narrow corollas, long pedicels, exerted stigmas.

%Yellow= no anthocyanin, long wide corollas, short pedicels, inserted stigmas.

%import data from raw floral traits, only ecotype, latitude and longitude

%as matrix

%red flowers are rows 1-82 hybrid are 83-130 yellow are 131-185

red\_location=RawfloraltraitdataS1(1:81,:,:);

hydrid\_location=RawfloraltraitdataS1(82:129,:,:);

yellow\_location=RawfloraltraitdataS1(130:184,:,:);

%scatter(RawfloraltraitdataS1)

%Plotting sampling locations wiht lattitude and longitude

hold on

scatter(red\_location(:,[3]), red\_location(:,[2]),'filled','diamond', 'r')

scatter(hydrid\_location(:,[3]), hydrid\_location(:,[2]),'filled','pentagram', 'k')

scatter(yellow\_location(:,[3]), yellow\_location(:,[2]),'filled','diamond', 'y')

title('Flower Sampling Locations in San Diego')

ylabel('Latitude')

xlabel('Longitude')

legend('Red', 'Hybrid', 'Yellow')

hold off

%%

Pedicel=FloraltraitsreducedS1(:,[6]) %selects only the pedicel variable

red\_pedicel=(Pedicel(1:81))

hybrid\_pedicel=(Pedicel(82:129))

yellow\_pedicel=(Pedicel(130:184))

boxplot(Floraltraitsreduced3S2, 'Labels',{'Red','Hybrid','Yellow'})% plot a matrix where column 1 is red pedicel height, 2 is hybrid and 3 is yellow

xlabel('Flower ecotype')

ylabel('Flower pedicel length')

title('Distribution of flower heights')

hold on

scatter(FloraltraitsreducedS1(:,[1]), FloraltraitsreducedS1(:,[6]))

%plots box plots with scattered points!!

hold off

figure;

corollalength=FloraltraitsreducedS2(:,[8]) %all corolla length

%plot distribution of corolla length across red hybrid and yellow

boxplot(floralcorollalength, 'Labels',{'Red','Hybrid','Yellow'})

xlabel('Flower ecotype')

ylabel('Flower corolla length')

title('Distribution of corolla lengths')

hold on

scatter(FloraltraitsreducedS2(:,[1]), FloraltraitsreducedS2(:,[8]))

hold off

figure;

%plot distribution of corolla width

corollawidth=FloraltraitsreducedS2(:,[9]) %all corolla width

%plot distribution of corolla length across red hybrid and yellow

boxplot(floralcorollawidth, 'Labels',{'Red','Hybrid','Yellow'})

xlabel('Flower ecotype')

ylabel('Flower corolla width')

title('Distribution of corolla widths')

hold on

scatter(FloraltraitsreducedS2(:,[1]), FloraltraitsreducedS2(:,[9]))

hold off

%% pigment

antho\_pigment=FloraltraitsreducedS2(:,[4]) %pigment absorbance at 520nm

%import matrix with just antho data in 3 columns

boxplot(anthopigment, 'Labels',{'Red','Hybrid','Yellow'})

xlabel('Flower ecotype')

ylabel('Anthocyanin absorbance')

title('Red pigmentation')

hold on

scatter(FloraltraitsreducedS2(:,[1]), FloraltraitsreducedS2(:,[4]))

hold off

figure;

%%

%Import floral traits reduced as matrix, I replaced NaNs with 0's

floraltraits\_transpose=FloraltraitsreducedS1' %transpose matrix now rows are variables

coeff = pca(FloraltraitsreducedS1) %rows are observations output is pxp of columns

[coefficient,score,latent] = pca(FloraltraitsreducedS1)

%coefficient is 11x11 score is 184x11 latent=11x1

Xcentered = score\*coefficient

%Each column of score corresponds to one principal component. The vector, latent, stores the variances of the four principal components. Reconstruct the centered

scatter3(score(:,1),score(:,2),score(:,3)) %plots the first 3 scores

%how much does each variable explain the data

[coeff2,score2,latent2,tsquared2,explained2] = pca(FloraltraitsreducedS1)

%explained gives how much variability is explained by a component

%% Correlations between observations of traits

%data in cross traits excel file, red

red\_correlations=corr(crosstraitsRed); %returns normalized correlations between all 6 traits

red\_c=heatmap(xvalues, yvalues,red\_correlations);

red\_c.Colormap = bone

red\_c.Title='Trait correaltions in red flowers';

figure;

yellow\_correlations=corr(crosstraitsYellow); %returns normalized correlations between all 6 traits

yellow\_c=heatmap(xvalues, yvalues,yellow\_correlations);

yellow\_c.Colormap = bone

yellow\_c.Title='Trait correaltions in yellow flowers';

figure;

F1\_correlations=corr(crosstraitsF1);

F1\_c=heatmap(xvalues, yvalues,F1\_correlations);

F1\_c.Colormap = bone

F1\_c.Title='Trait correaltions in F1 cross';

figure;

F2\_correlations=corr(crosstraitsF2);

F2\_c=heatmap(xvalues, yvalues, F2\_correlations); %creates a heatmap of the normalized correlations and labels the axis

F2\_c.Colormap = bone

xvalues = {'tube width','petal length', 'petal width','anthocyanin absorbance','stem length','stigma exertion'}

yvalues = {'tube width','petal length', 'petal width','anthocyanin absorbance','stem length','stigma exertion'}

F2\_c.Title='Trait correaltions in F2 cross';

%% PCA

[coeffRed, pcRed,~,~,explainedRed] = pca(crosstraitsRed)

pcRed=pcRed';

coeffRed=coeffRed';

scatter(pcRed(1,:),pcRed(2,:),'filled','r'); %plots PC1 and PC2

hold on

[coeffYellow, pcYellow,~,~,explainedYellow] = pca(crosstraitsYellow) ;

pcYellow=pcYellow';

coeffYellow=coeffYellow';

scatter(pcYellow(1,:),pcYellow(2,:),'filled','yellow'); %plots PC1 and PC2hold off

[coeffF1, pcF1,~,~,explainedF1] = pca(crosstraitsF1);

%W is the principal component coefficients in pxp matrix, pc are the scores

%Principal component scores are the representations of X in the principal component space.

pcF1=pcF1';

coeffF1=coeffF1';

scatter(pcF1(1,:),pcF1(2,:),'o','magenta'); %plots PC1 and PC2

% title('{\bf PCA} by princomp'); xlabel('PC 1'); ylabel('PC 2')

[coeffF2, pcF2,~,~,explained] = pca(crosstraitsF2) ;

%W is the principal component coefficients in pxp matrix, pc are the scores

%Principal component scores are the representations of X in the principal component space.

pcF2=pcF2';

coeffF2=coeffF2';

scatter(pcF2(1,:),pcF2(2,:),'x','black'); %plots PC1 and PC2

title('{\bf PCA}'); xlabel('PC 1 (63.45%)'); ylabel('PC 2 (22.45%)')

legend('Red', 'Yellow', 'F1', 'F2')

hold off

figure;

%PCA of all the data together

[coeffall, pcall,~,~,explainedall] = pca(crosstraits) ;

scatter(pcall(:,1),pcall(:,2));

coeffF1