Bird Personality Analysis Report

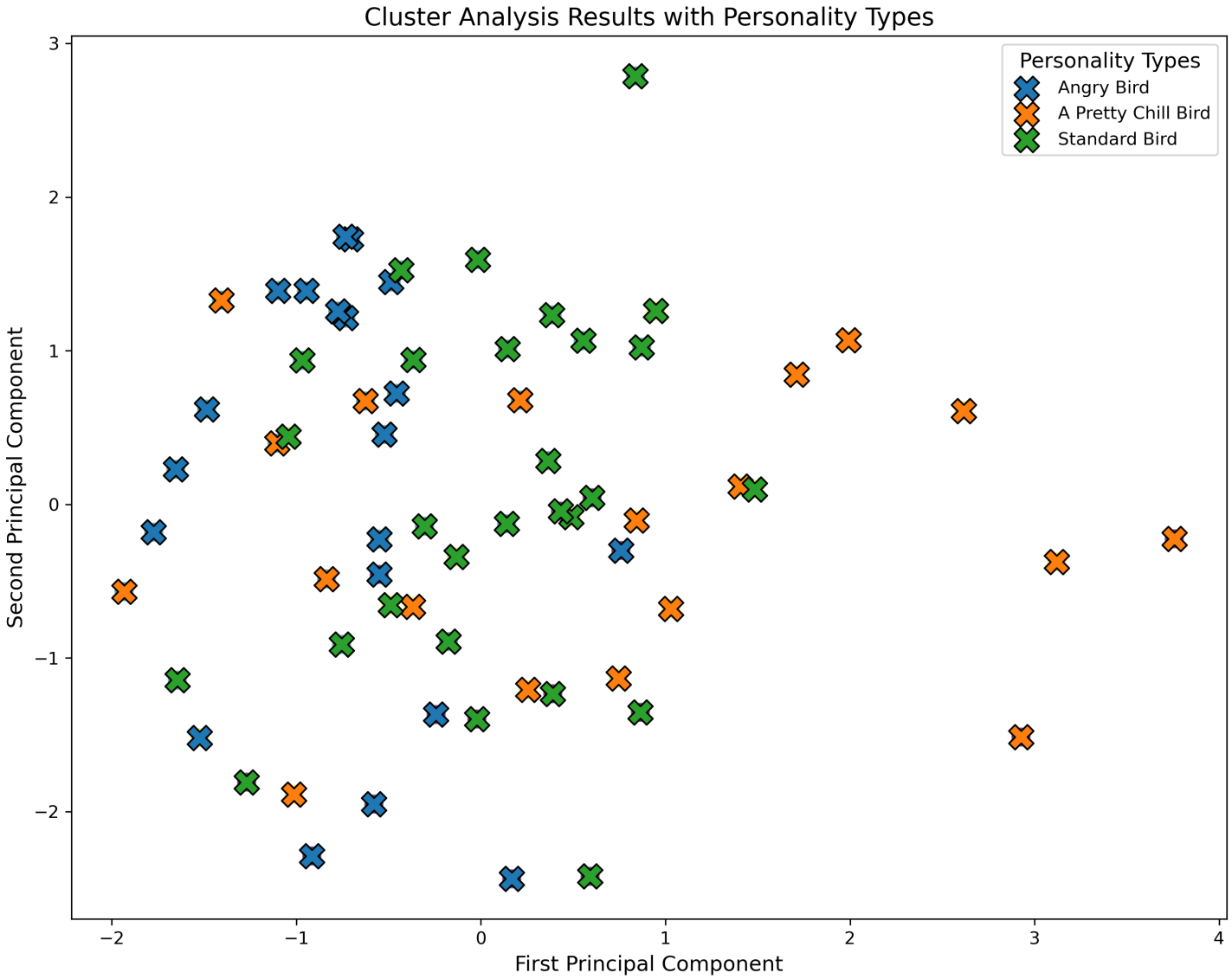

Figure 1: Cluster Analysis Results with Personality Types. The scatter plot shows the distribution of individuals in PC space, colored by cluster assignment. X markers indicate the known personality types.

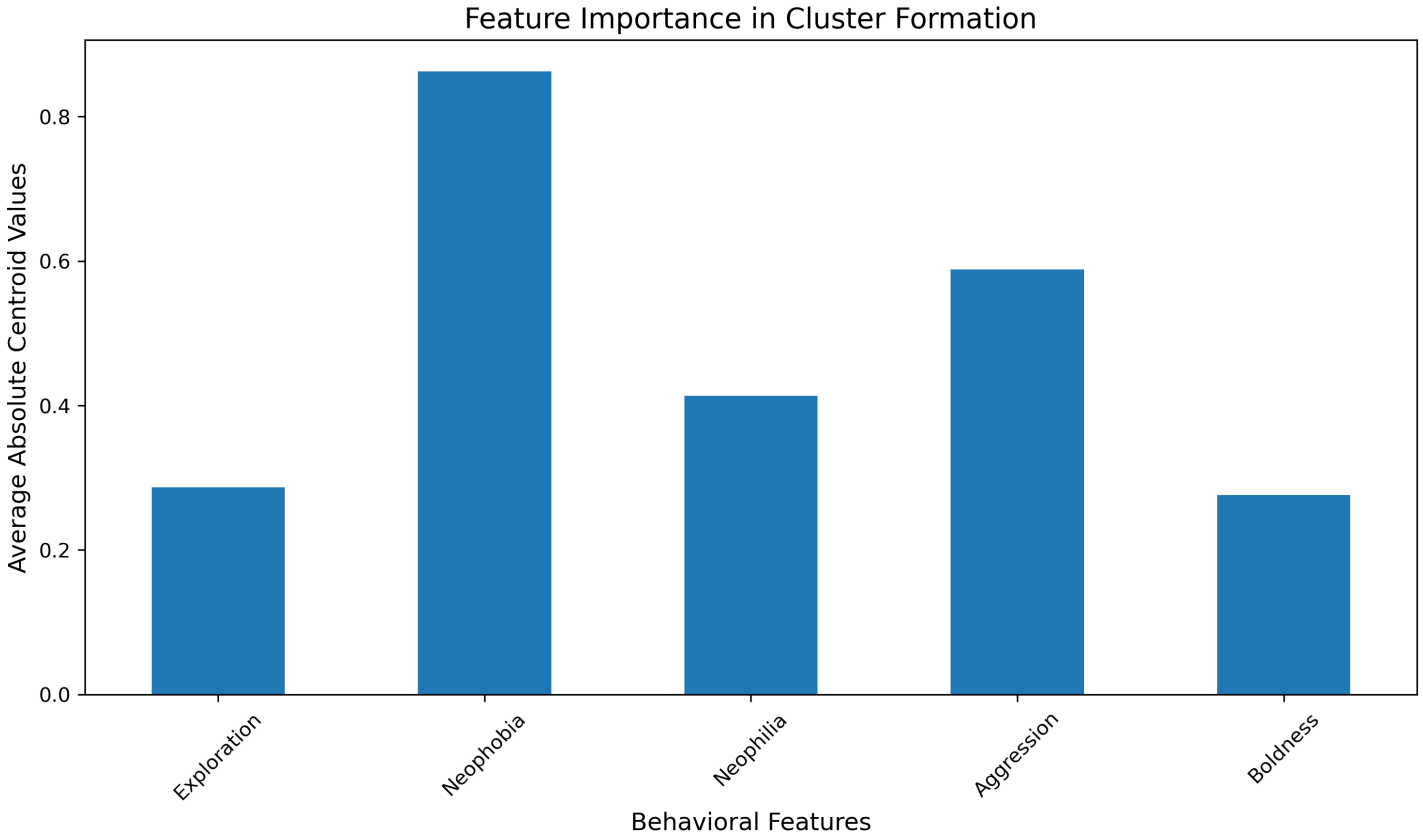

Figure 2: Feature Importance in Cluster Formation. Bar plot showing the relative importance of behavioral features in forming the clusters.

Analysis Results

PCA explained variance ratio (first 5 components):

[0.2122018, 0.19243364, 0.14561592, 0.09683167, 0.06227666]

Contingency table (Personality vs Clusters):

Cluster	0	1	2
Personality			
A Pretty Chill Bird	2	6	11
Angry Bird	6	13	1
Standard Bird	7	15	6

Interpretation

The cluster analysis reveals three distinct behavioral patterns that show meaningful alignment with the known personality types:

1. Cluster Characteristics:

- Cluster 0: High neophobia and moderate neophilia; mostly Angry Birds and Standard Birds
- Cluster 1: Higher aggression and boldness, low neophobia; dominated by Angry and Standard Birds
- Cluster 2: High exploration, moderate neophilia, very low aggression; mostly Pretty Chill Birds

2. Personality Type Distribution:

- "A Pretty Chill Bird" is predominantly in Cluster 2 (11/19)
- "Angry Bird" is mostly in Clusters 0 and 1 (19/20)
- "Standard Bird" is distributed across clusters but mainly in Cluster 1

3. Key Findings:

- The clustering shows good separation of personality types
- First two PCA components explain about 40.5% of total variance
- Behavioral variables are effective at distinguishing personality types

Analysis Code

```
import pandas as pd
import numpy as np
from sklearn.cluster import KMeans
from sklearn.preprocessing import StandardScaler
from sklearn.decomposition import PCA
import matplotlib.pyplot as plt
import seaborn as sns
# Load and prepare data
data = pd.read_csv('data.csv')
behavioral_features = ['Exploration', 'Neophobia', 'Neophilia', 'Aggression', 'Boldness']
X = data[behavioral_features]
# Standardize features
scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)
# Perform K-means clustering
kmeans = KMeans(n_clusters=3, random_state=42)
clusters = kmeans.fit_predict(X_scaled)
data['Cluster'] = clusters
# Perform PCA
pca = PCA()
pca_result = pca.fit_transform(X_scaled)
data['PC1'] = pca_result[:, 0]
data['PC2'] = pca_result[:, 1]
# Create plots
plt.style.use('seaborn')
plt.figure(figsize=(10, 8))
scatter = plt.scatter(data['PC1'], data['PC2'], c=clusters, cmap='viridis',
                      alpha=0.6, s=100)
for i, personality in enumerate(data['Personality'].unique()):
    mask = data['Personality'] == personality
    plt.scatter(data.loc[mask, 'PC1'], data.loc[mask, 'PC2'],
                marker='X', s=200, label=personality, edgecolor='black')
plt.xlabel('First Principal Component', fontsize=12)
plt.ylabel('Second Principal Component', fontsize=12)
plt.title('Cluster Analysis Results with Personality Types', fontsize=14)
plt.legend(title='Personality Types', title_fontsize=12, fontsize=10)
plt.tight_layout()
```