In [1]:
```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt

# Load the dataset
file_path = '/Users/pavanijain/Desktop/2013_nba_draft_combine.csv'
data = pd.read_csv(file_path)

# Display the first few rows of the dataset
print("First few rows of the dataset:")
print(data.head())
```

```
First few rows of the dataset:
   Unnamed: 0          Player  Year  Draft pick  Height (No Shoes)  \
0           0   Adonis Thomas  2013         NaN              76.75
1           1    Allen Crabbe  2013        31.0              77.25
2           2  Andre Roberson  2013        26.0              78.25
3           3  Archie Goodwin  2013        29.0              75.75
4           4      B.J. Young  2013         NaN              74.25

   Height (With Shoes)  Wingspan  Standing reach  Vertical (Max)  \
0                77.75     85.00            99.0            40.5
1                78.25     83.25           103.5            36.0
2                79.00     83.00           104.5            36.5
3                77.25     81.50           102.0            36.0
4                75.50     80.25            99.0             NaN

   Vertical (Max Reach)  Vertical (No Step)  Vertical (No Step Reach)  Weig
ht  \
0                 139.5                34.5                     133.5   23
2.0
1                 139.5                30.5                     134.0   19
7.0
2                 141.0                30.0                     134.5   20
6.0
3                 138.0                30.0                     132.0   18
9.0
4                   NaN                 NaN                       NaN   17
9.0

   Body Fat  Hand (Length)  Hand (Width)  Bench  Agility  Sprint
0       7.5           9.25           9.5   13.0    11.66    3.32
1       4.7           8.25           8.5   10.0    10.67    3.32
2       7.1           8.75           9.5   10.0    11.36    3.34
3       4.6           8.50           9.0    1.0    10.75    3.27
4       4.7           8.50           9.5    NaN      NaN     NaN
```

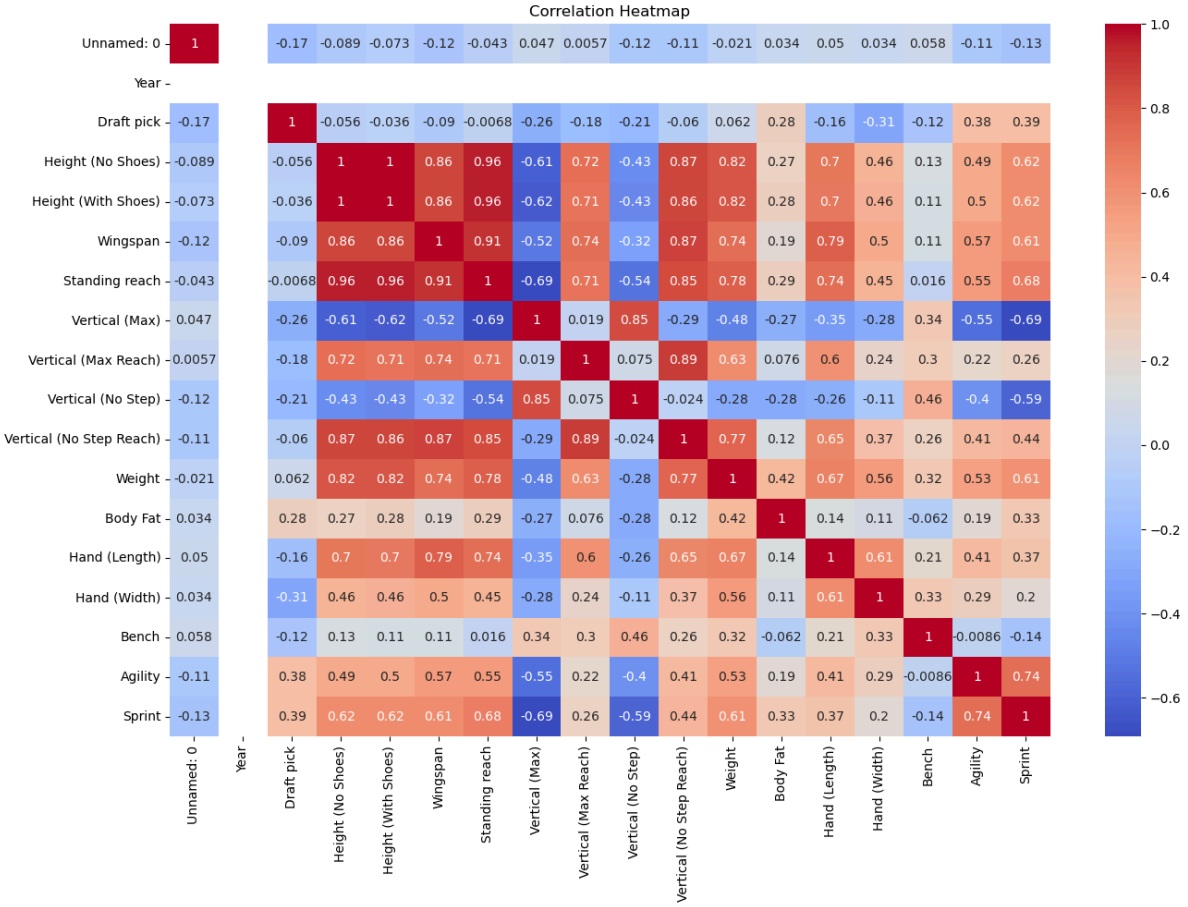
In [2]:
```
# Display basic information about the dataset
print("\nBasic Information:")
print(data.info())

# Summary statistics
print("\nSummary Statistics:")
print(data.describe())

# Checking for missing values
print("\nMissing Values:")
print(data.isnull().sum())
```

```
Basic Information:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 62 entries, 0 to 61
Data columns (total 19 columns):
 #   Column                    Non-Null Count  Dtype  
---  ------                    --------------  -----  
 0   Unnamed: 0                62 non-null     int64  
 1   Player                    62 non-null     object 
 2   Year                      62 non-null     int64  
 3   Draft pick                44 non-null     float64
 4   Height (No Shoes)         62 non-null     float64
 5   Height (With Shoes)       62 non-null     float64
 6   Wingspan                  62 non-null     float64
 7   Standing reach            62 non-null     float64
 8   Vertical (Max)            51 non-null     float64
 9   Vertical (Max Reach)      51 non-null     float64
 10  Vertical (No Step)        51 non-null     float64
 11  Vertical (No Step Reach)  51 non-null     float64
 12  Weight                    62 non-null     float64
 13  Body Fat                  62 non-null     float64
 14  Hand (Length)             62 non-null     float64
 15  Hand (Width)              62 non-null     float64
 16  Bench                     52 non-null     float64
 17  Agility                   51 non-null     float64
 18  Sprint                    51 non-null     float64
dtypes: float64(16), int64(2), object(1)
memory usage: 9.3+ KB
None

Summary Statistics:
       Unnamed: 0    Year  Draft pick  Height (No Shoes)  Height (With Shoe
s)  \
count   62.000000    62.0   44.000000          62.000000            62.0000
00
mean    30.500000  2013.0   28.545455          77.471774            78.8024
19
std     18.041619     0.0   16.192813           3.729618             3.7411
47
min      0.000000  2013.0    2.000000          69.500000            70.5000
00
25%     15.250000  2013.0   13.750000          75.500000            76.7500
00
50%     30.500000  2013.0   28.000000          77.375000            79.0000
00
75%     45.750000  2013.0   41.250000          80.687500            81.8750
00
max     61.000000  2013.0   58.000000          84.500000            86.0000
00

        Wingspan  Standing reach  Vertical (Max)  Vertical (Max Reach)  \
count  62.000000       62.000000       51.000000             51.000000   
mean   82.000000      103.104839       36.186275            139.313725   
std     4.405427        5.440541        3.595776              3.724192   
min    70.000000       89.500000       29.000000            130.500000   
25%    79.250000       99.250000       33.750000            137.000000   
50%    82.375000      103.750000       36.500000            139.500000   
75%    85.500000      106.875000       38.500000            142.000000   
max    92.500000      115.000000       44.000000            145.000000   

       Vertical (No Step)  Vertical (No Step Reach)      Weight   Body Fat
\
count           51.000000                 51.000000   62.000000  62.000000
mean            30.107843                133.235294  209.677419   6.574194
std              2.663294                  4.317815   23.547492   2.072453
```

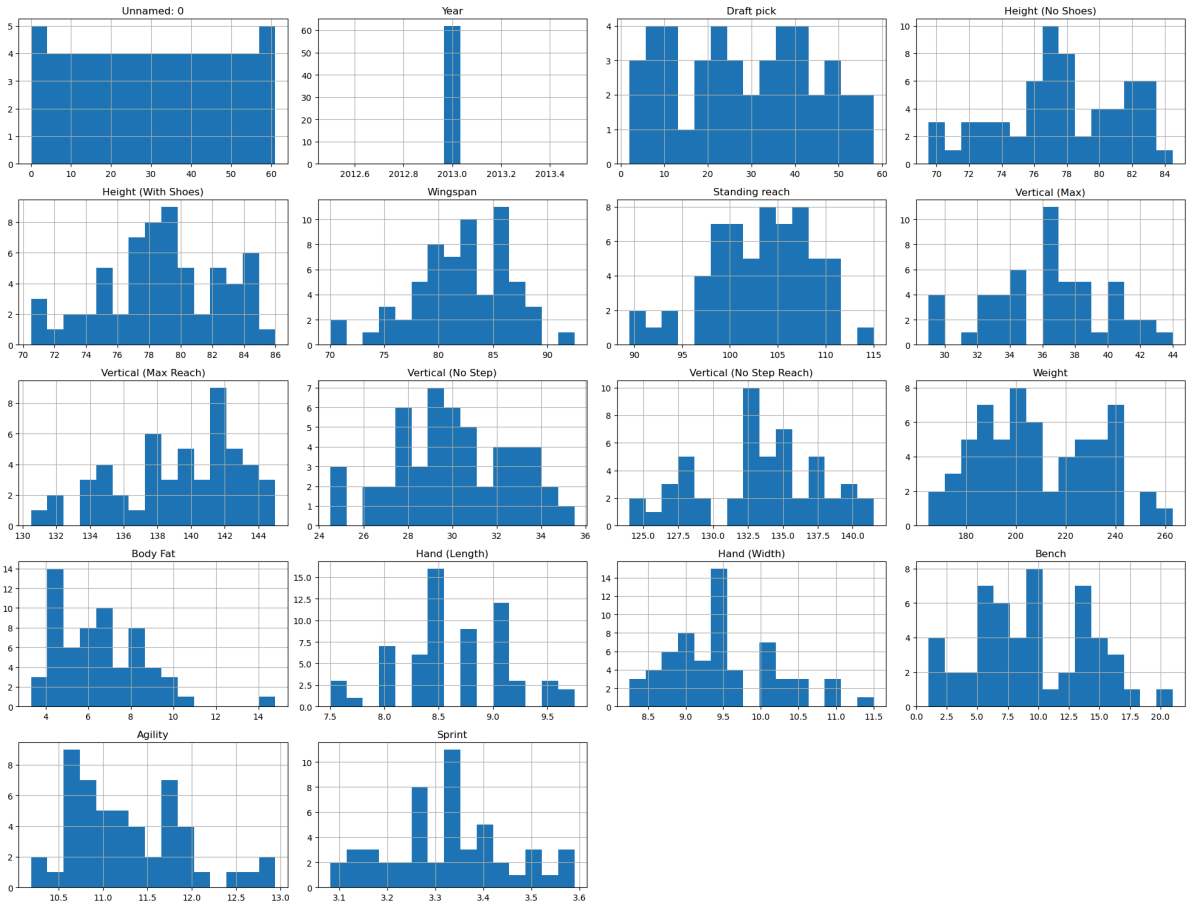
```
min             24.500000                124.000000  165.000000   3.300000
25%             28.250000                130.500000  191.000000   4.725000
50%             30.000000                133.500000  206.000000   6.350000
75%             32.250000                136.250000  230.000000   8.000000
max             35.500000                141.500000  263.000000  14.800000

       Hand (Length)  Hand (Width)      Bench    Agility     Sprint
count      62.000000     62.000000  52.000000  51.000000  51.000000
mean        8.616935      9.479839   9.250000  11.281765   3.321569
std         0.515573      0.721875   4.714912   0.649491   0.128411
min         7.500000      8.250000   1.000000  10.190000   3.080000
25%         8.250000      9.000000   5.000000  10.760000   3.250000
50%         8.500000      9.500000   9.000000  11.160000   3.320000
75%         9.000000     10.000000  13.000000  11.725000   3.400000
max         9.750000     11.500000  21.000000  12.940000   3.590000

Missing Values:
Unnamed: 0                   0
Player                       0
Year                         0
Draft pick                  18
Height (No Shoes)            0
Height (With Shoes)          0
Wingspan                     0
Standing reach               0
Vertical (Max)              11
Vertical (Max Reach)        11
Vertical (No Step)          11
Vertical (No Step Reach)    11
Weight                       0
Body Fat                     0
Hand (Length)                0
Hand (Width)                 0
Bench                       10
Agility                     11
Sprint                      11
dtype: int64
```

In [3]:

```
# Visualizations
# Correlation heatmap (only for numeric columns)
numeric_data = data.select_dtypes(include=[np.number])
plt.figure(figsize=(14, 10))
sns.heatmap(numeric_data.corr(), annot=True, cmap='coolwarm')
plt.title('Correlation Heatmap')
plt.tight_layout()
plt.show()
```

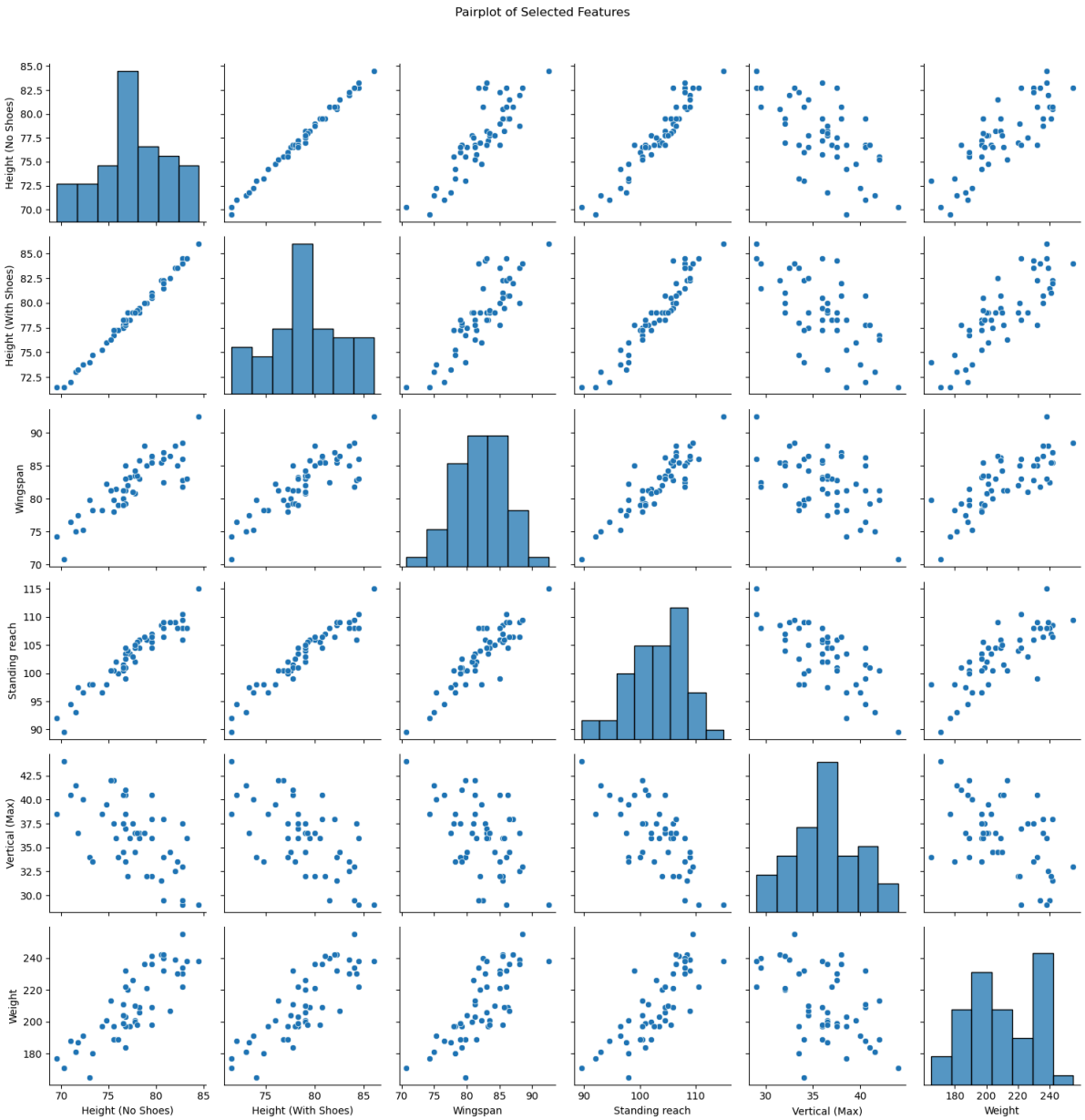


In [4]:
```
# Distribution of numeric features
numeric_data.hist(bins=15, figsize=(20, 15), layout=(5, 4))
plt.tight_layout()
plt.show()
```

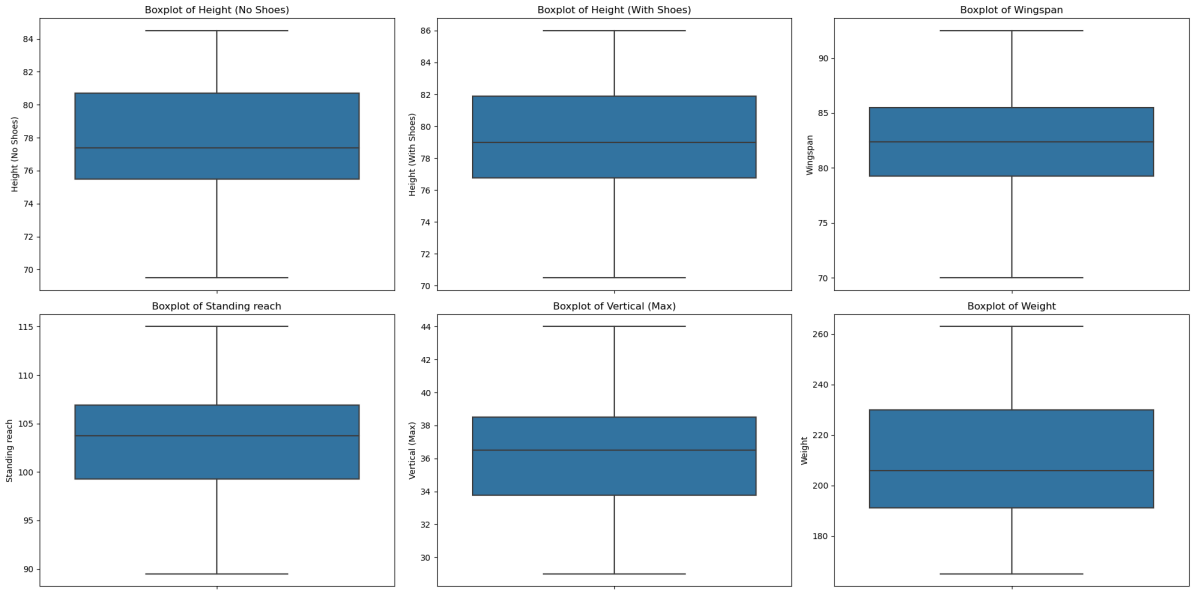


In [10]:
```
# Pairplot of selected features
selected_features = ['Height (No Shoes)', 'Height (With Shoes)', 'Wingspan',
sns.pairplot(data[selected_features].dropna())
plt.suptitle('Pairplot of Selected Features', y=1.02)
plt.tight_layout()
plt.show()
```

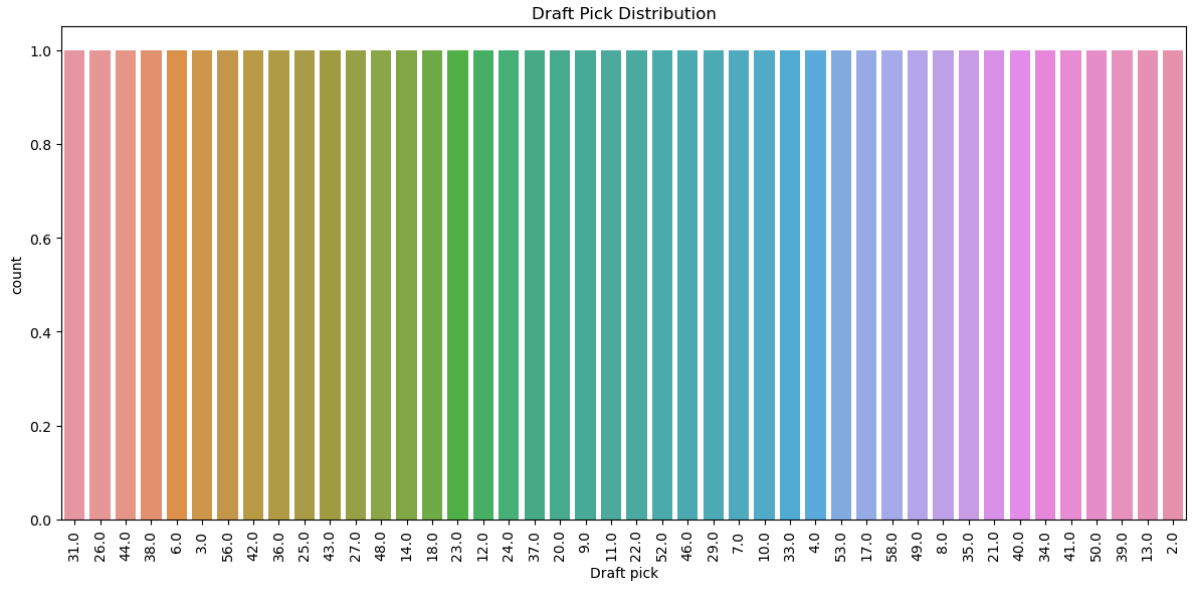
```
/Users/pavanijain/anaconda3/lib/python3.11/site-packages/seaborn/axisgrid.p
y:118: UserWarning: The figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
/var/folders/vw/vk20m1ys0tz70yl7_xyhtqvc0000gn/T/ipykernel_6544/680349118.p
y:5: UserWarning: The figure layout has changed to tight
  plt.tight_layout()
```



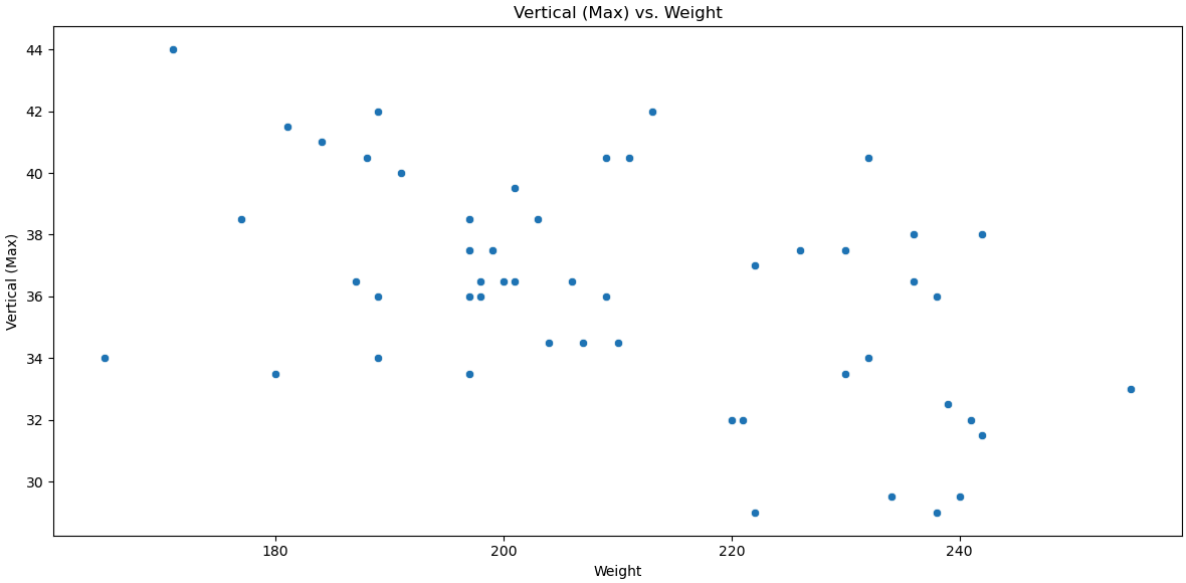
In [11]:
```
# Boxplots of selected features
plt.figure(figsize=(20, 10))
for i, feature in enumerate(selected_features):
    plt.subplot(2, 3, i+1)
    sns.boxplot(y=data[feature])
    plt.title(f'Boxplot of {feature}')
plt.tight_layout()
plt.show()
```



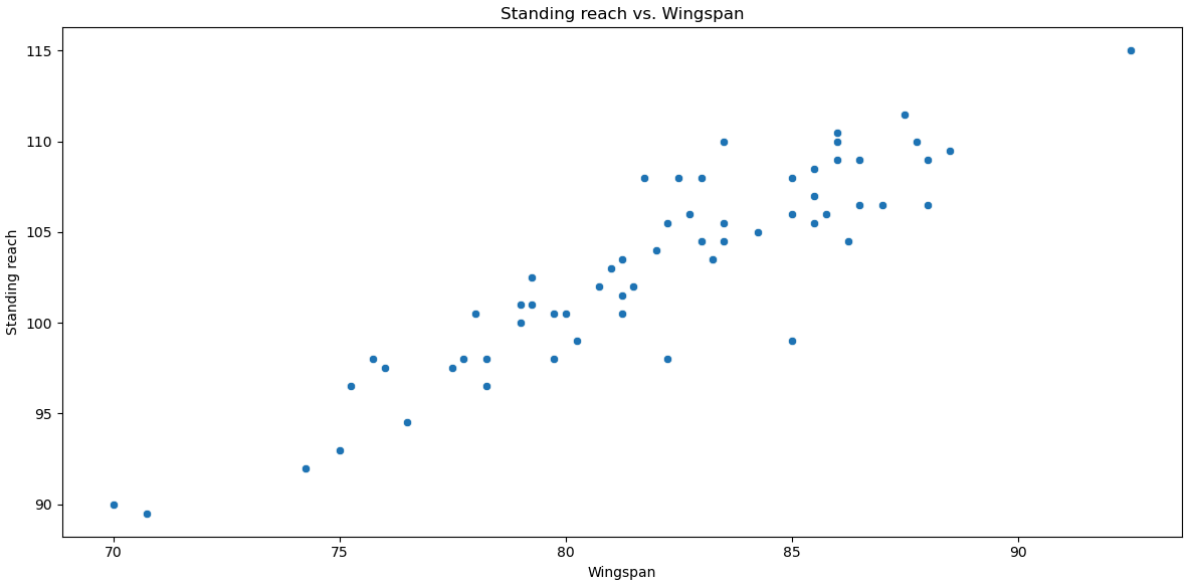
In [12]:
```
# Draft pick distribution
plt.figure(figsize=(12, 6))
sns.countplot(x='Draft pick', data=data, order=data['Draft pick'].dropna().v
plt.title('Draft Pick Distribution')
plt.xticks(rotation=90)
plt.tight_layout()
plt.show()
```



In [13]:
```
# Vertical (Max) vs. Weight scatter plot
plt.figure(figsize=(12, 6))
sns.scatterplot(x='Weight', y='Vertical (Max)', data=data)
plt.title('Vertical (Max) vs. Weight')
plt.tight_layout()
plt.show()
```



In [14]:
```
# Standing reach vs. Wingspan scatter plot
plt.figure(figsize=(12, 6))
sns.scatterplot(x='Wingspan', y='Standing reach', data=data)
plt.title('Standing reach vs. Wingspan')
plt.tight_layout()
plt.show()
```



In [ ]: