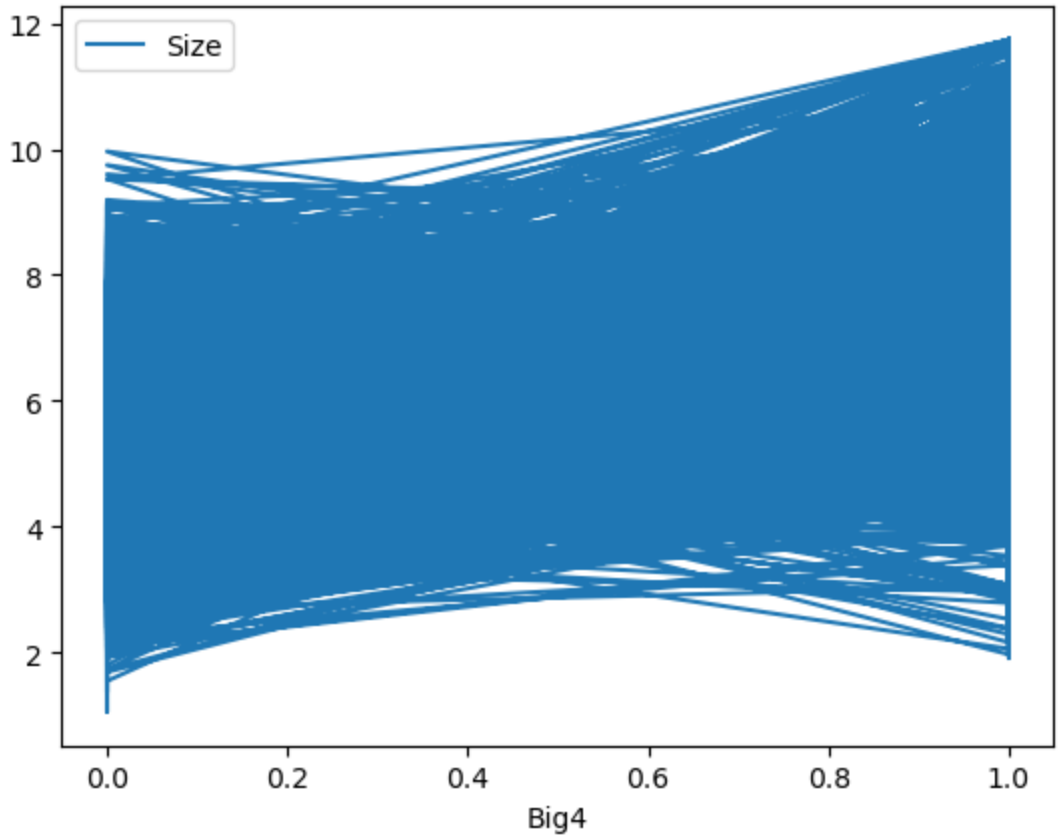
In [ ]:
```
# import pandas as pd, numpy as np
pd.set_option('display.width',150)
pd.set_option('display.max_columns',20)
df = pd.read_csv("GCO_Sample_Data_2022.csv")
df.columns
```

In [5]:
```
df.plot(x='Big4',y='Size')
```

Out[5]:
```
<Axes: xlabel='Big4'>
```



In [6]:
```
col1, col2 = "Size", "Big4"
corr = df[col1].corr(df[col2])
print ("Correlation between ", col1, " and ", col2, "is: ", round(corr, 2)) #
```

```
Correlation between  Size  and  Big4 is:  0.41
```

In [7]:
```
df.groupby(['GCO']).agg({'Big4':'mean','Loss':lambda x: list(x)})
```

Out[7]:

GCO	Big4	Loss
0	0.780877	[0, 1, 1, 0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 1, 1, ...
1	0.566038	[1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1, ...

In [8]:
```
df.isnull().sum() # Return null values
df['GCO'].value_counts()
```

Out[8]:
```
0    4947
1      53
Name: GCO, dtype: int64
```

In [9]:
```
# Replace according to the mapping table provided above
df['goingconcern'] = df['GCO'].replace({1 : 'GC_Warning', 0: 'Opinion_Clean'})
df.goingconcern.value_counts()
```

Out[9]:
```
Opinion_Clean    4947
GC_Warning         53
Name: goingconcern, dtype: int64
```

In [10]:
```
df['GCO'].describe()
```

Out[10]:
```
count    5000.000000
mean        0.010600
std         0.102419
min         0.000000
25%         0.000000
50%         0.000000
75%         0.000000
max         1.000000
Name: GCO, dtype: float64
```

In [11]:
```
# Approach 1 - Using Pandas
print(f"Using Pandas:")
print(f"mean: {df.GCO.mean()}")
print(f"standard_deviation: {df.GCO.std()}")
print(f"minimum: {df.GCO.min()}")
print(f"25th_percentile: {df.GCO.quantile(0.25)}")
print(f"50th_percentile: {df.GCO.quantile(0.50)}")
print(f"75th_percentile: {df.GCO.quantile(0.75)}")
print(f"maximum: {df.GCO.max()}\n")
```

```
Using Pandas:
mean: 0.0106
standard_deviation: 0.10241942173040403
minimum: 0
25th_percentile: 0.0
50th_percentile: 0.0
75th_percentile: 0.0
maximum: 1
```

In [12]:
```
# Approach 2 - Using NumPy
print(f"Using NumPy:")
print(f"mean: {np.mean(df.GCO)}")
print(f"standard_deviation: {np.std(df.GCO, ddof = 1)}")
print(f"minimum: {np.min(df.GCO)}")
print(f"25th_percentile: {np.percentile(df.GCO, 25)}")
print(f"50th_percentile: {np.percentile(df.GCO, 50)}")
print(f"75th_percentile: {np.percentile(df.GCO, 75)}")
print(f"maximum: {np.max(df.GCO)}\n")
```

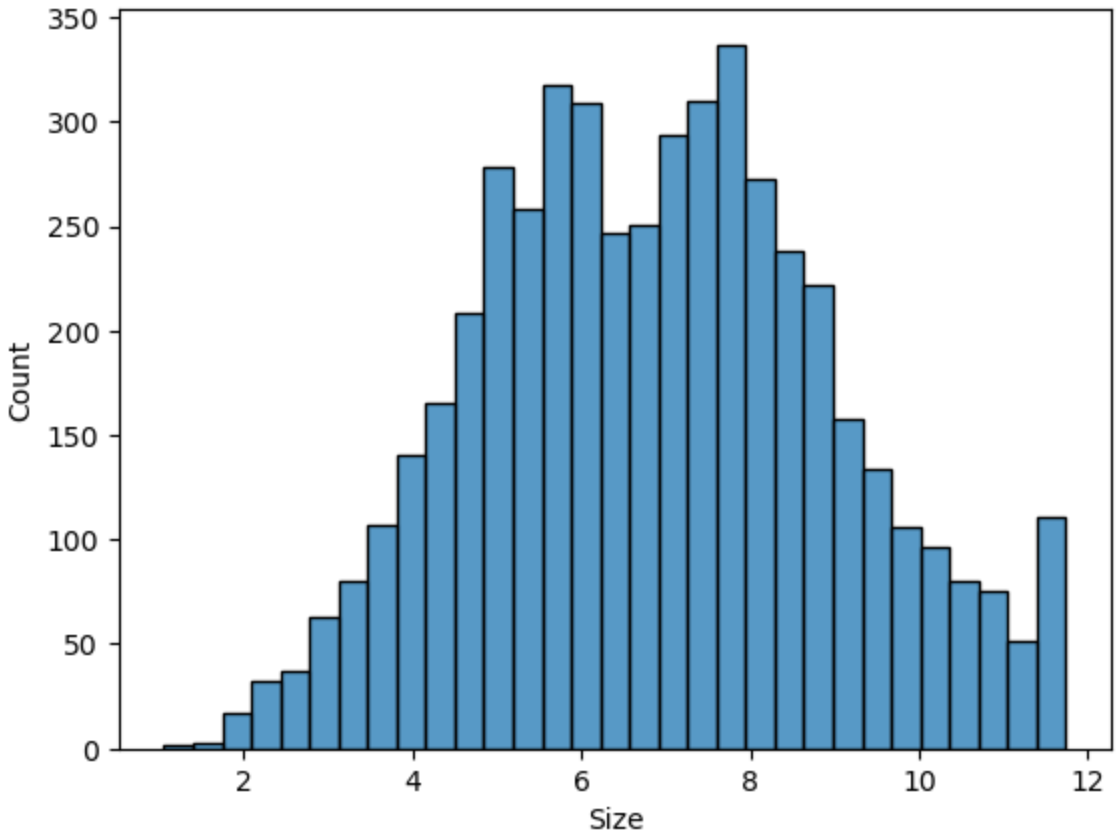
```
Using NumPy:
mean: 0.0106
standard_deviation: 0.10241942173040403
minimum: 0
25th_percentile: 0.0
50th_percentile: 0.0
75th_percentile: 0.0
maximum: 1
```

In [13]:
```
lower_bound = np.mean(df['Size'][df.GCO == 1])
upper_bound = np.mean(df['Size'][df.GCO == 0])
```

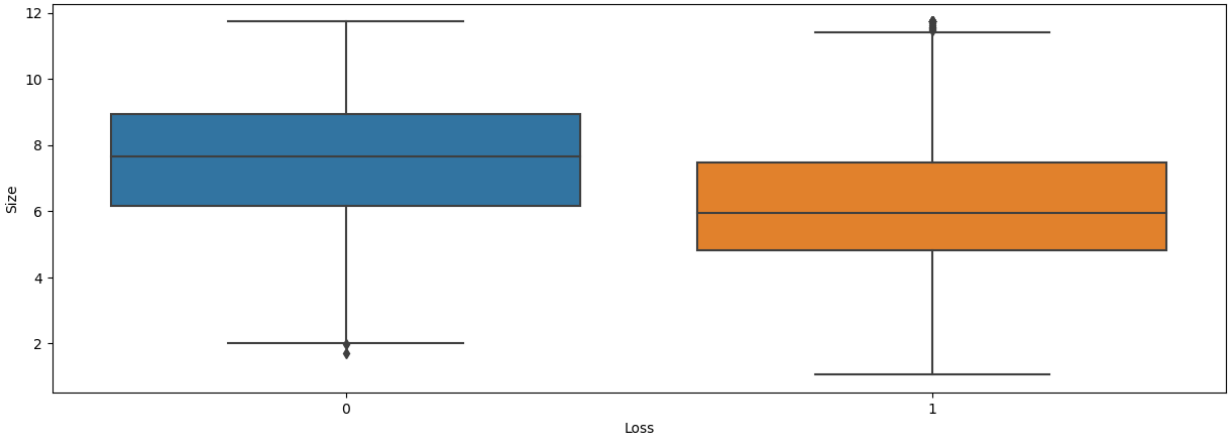
In [14]:
```
print(f"lower: {lower_bound}")
print(f"upper: {upper_bound}")
```

```
lower: 4.109821029433963
upper: 6.928418017709723
```

In [15]:
```
import seaborn as sns, matplotlib.pyplot as plt
sns.histplot(df.Size)
plt.show()
```



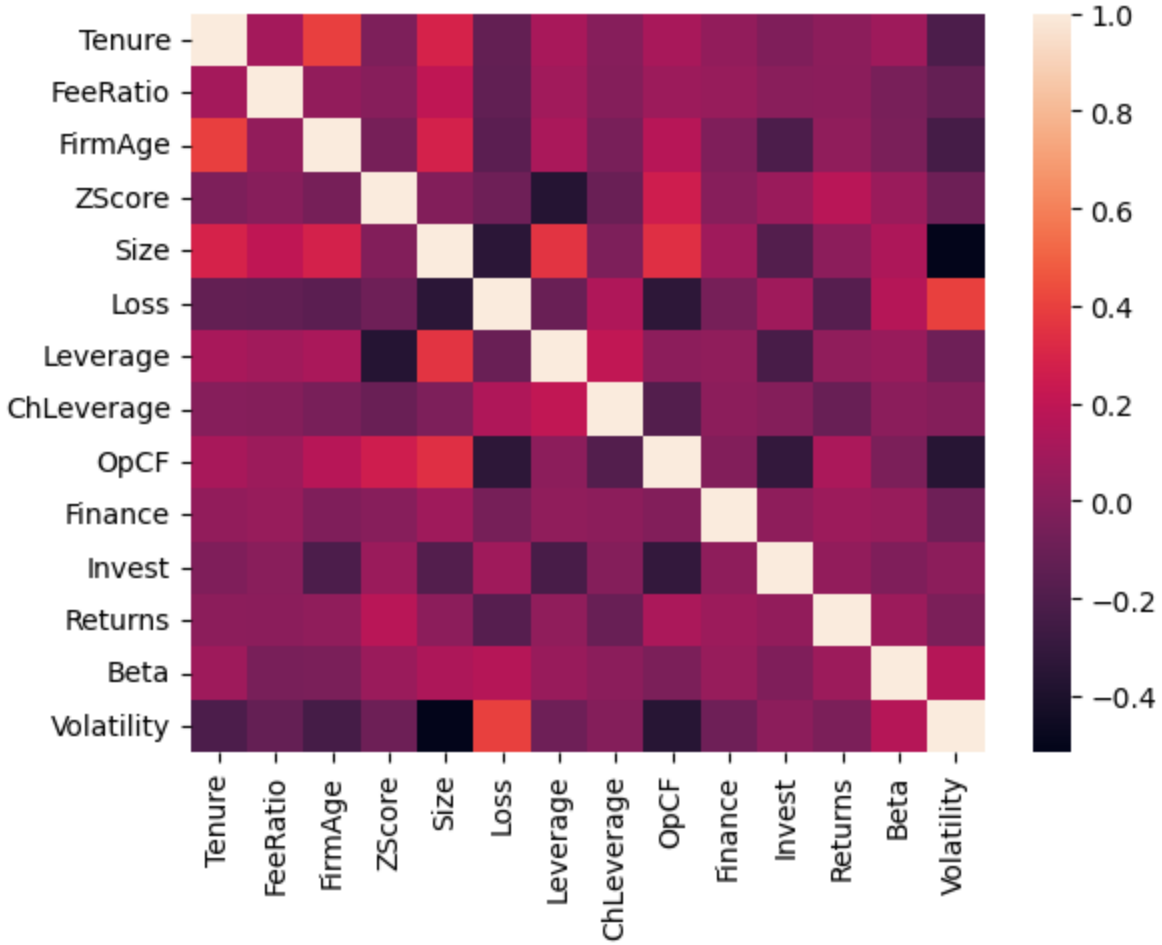
In [16]:
```
plt.figure(figsize = (15, 5))
sns.boxplot(data = df, x = 'Loss', y = 'Size')
plt.show()
```



In [19]:
```
xvars = ['Tenure','FeeRatio','FirmAge','ZScore','Size','Loss','Leverage',
         'ChLeverage','OpCF','Finance','Invest','Returns','Beta','Volatility']
# train an XGBoost model
X, y = df[xvars],df['Big4']
```

In [20]:
```
sns.heatmap(df[xvars].corr())
```

Out[20]:
```
<Axes: >
```



In [23]:
```
pip install scikit-learn
```

```
Collecting scikit-learn
  Downloading scikit_learn-1.3.2-cp310-cp310-macosx_10_9_x86_64.whl (10.2 MB)
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 10.2/10.2 MB 21.0 MB/s eta 0:00:
0000:010:01
Collecting scipy>=1.5.0
  Downloading scipy-1.11.3-cp310-cp310-macosx_10_9_x86_64.whl (37.3 MB)
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 37.3/37.3 MB 29.2 MB/s eta 0:00:
0000:0100:01
Requirement already satisfied: joblib>=1.1.1 in ./anaconda3/lib/python3.10/sit
e-packages (from scikit-learn) (1.1.1)
Requirement already satisfied: numpy<2.0,>=1.17.3 in ./anaconda3/lib/python3.1
0/site-packages (from scikit-learn) (1.23.5)
Requirement already satisfied: threadpoolctl>=2.0.0 in ./anaconda3/lib/python
3.10/site-packages (from scikit-learn) (2.2.0)
Installing collected packages: scipy, scikit-learn
Successfully installed scikit-learn-1.3.2 scipy-1.11.3
Note: you may need to restart the kernel to use updated packages.
```

In [29]:
```
pip install xgboost
```

```
Collecting xgboost
  Downloading xgboost-2.0.1-py3-none-macosx_10_15_x86_64.macosx_11_0_x86_64.ma
cosx_12_0_x86_64.whl (2.2 MB)
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 2.2/2.2 MB 6.7 MB/s eta 0:00:00a
0:00:01
Requirement already satisfied: scipy in ./anaconda3/lib/python3.10/site-packag
es (from xgboost) (1.11.3)
Requirement already satisfied: numpy in ./anaconda3/lib/python3.10/site-packag
es (from xgboost) (1.23.5)
Installing collected packages: xgboost
Successfully installed xgboost-2.0.1
Note: you may need to restart the kernel to use updated packages.
```

In [30]:
```
from sklearn.model_selection import train_test_split
import xgboost as xgb
```

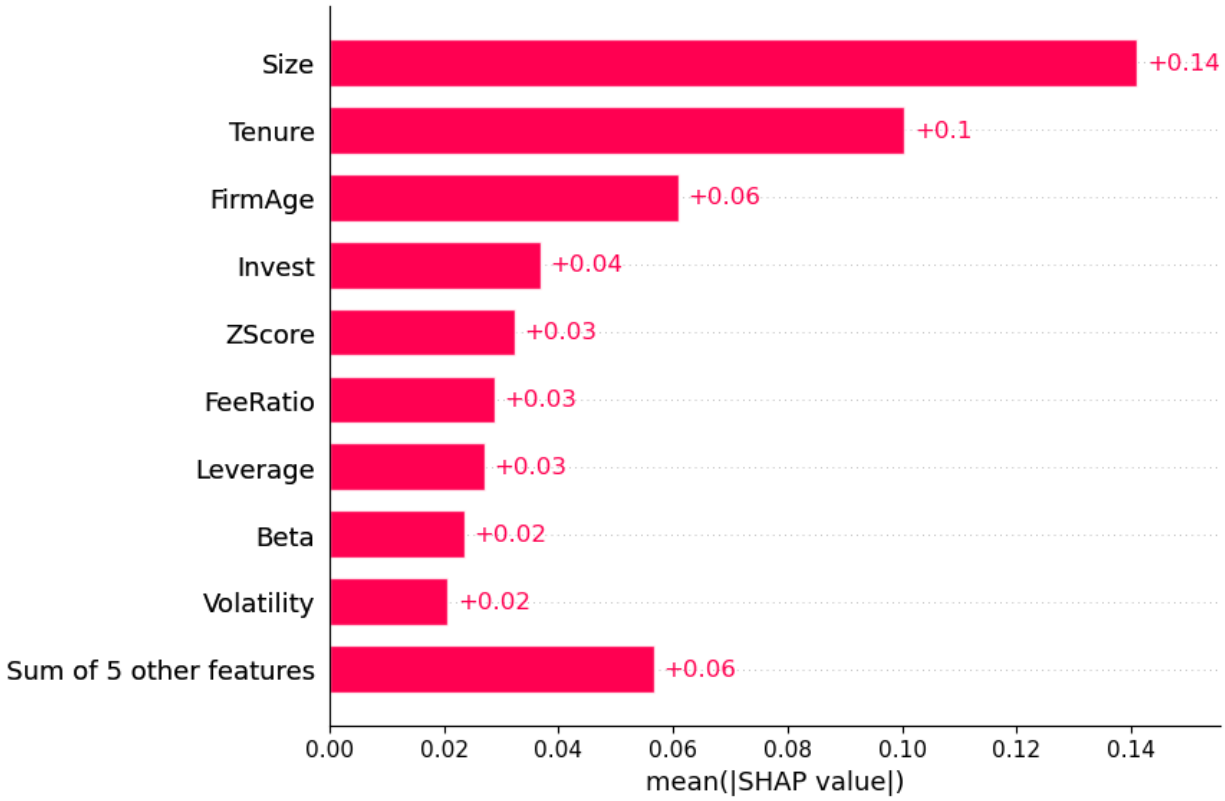
In [31]:
```
Y = df['Big4']
X =  df[xvars]
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size = 0.2, ran
xgb_model = xgb.XGBRegressor(random_state=42)
xgb_model.fit(X_train, Y_train)
```

Out[31]:
```
XGBRegressor(base_score=None, booster=None, callbacks=None,
             colsample_bylevel=None, colsample_bynode=None,
             colsample_bytree=None, device=None, early_stopping_rounds
=None,
             enable_categorical=False, eval_metric=None, feature_types
=None,
             gamma=None, grow_policy=None, importance_type=None,
             interaction_constraints=None, learning_rate=None, max_bin
=None,
             max_cat_threshold=None, max_cat_to_onehot=None,
             max_delta_step=None, max_depth=None, max_leaves=None,
```

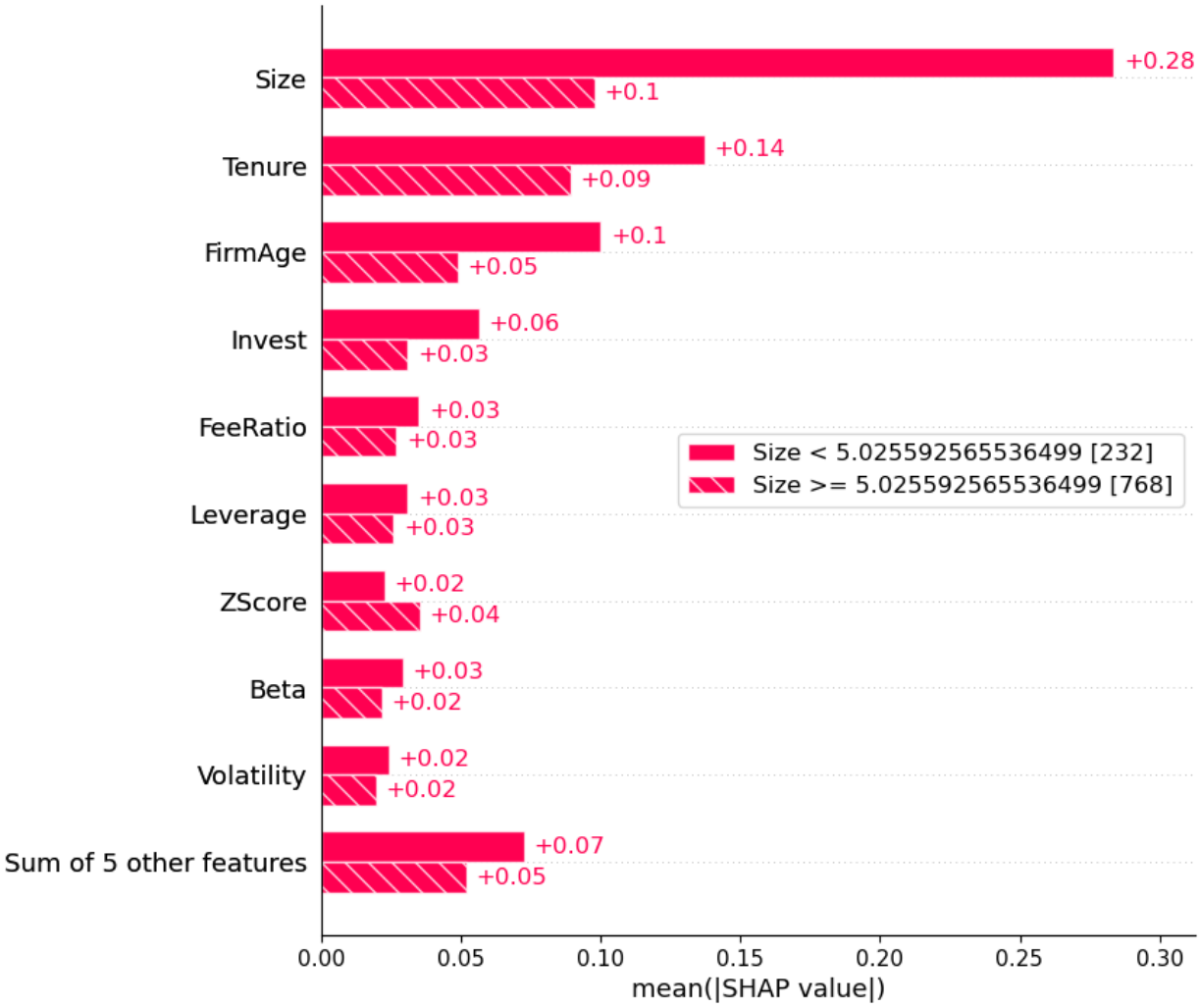
In [56]:
```
# The SHAP Values
import shap
```

```
explainer = shap.Explainer(xgb_model)
shap_values = explainer(X_test)
shap.plots.bar(shap_values, max_display=10) # default is max_display=12
```

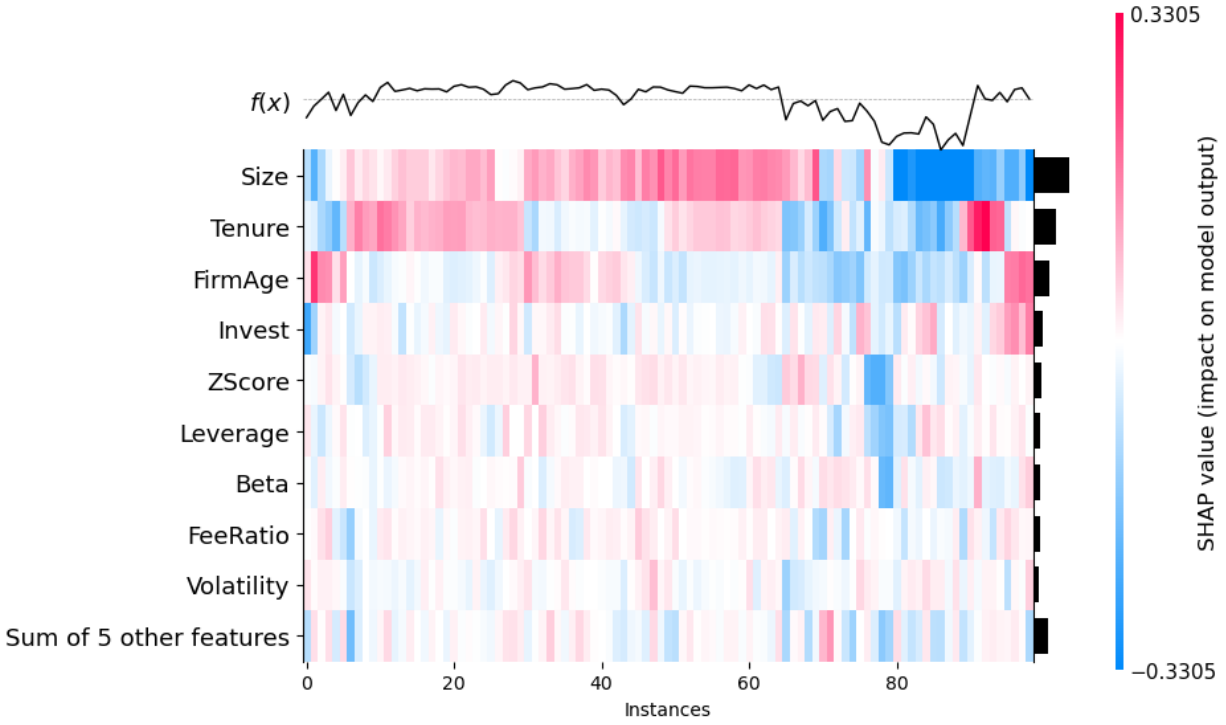
```
[17:24:22] WARNING: /Users/runner/work/xgboost/xgboost/src/c_api/c_api.cc:124
0: Saving into deprecated binary model format, please consider using `json` or
`ubj`. Model format will default to JSON in XGBoost 2.2 if not specified.
[17:24:23] WARNING: /Users/runner/work/xgboost/xgboost/src/c_api/c_api.cc:124
0: Saving into deprecated binary model format, please consider using `json` or
`ubj`. Model format will default to JSON in XGBoost 2.2 if not specified.
```



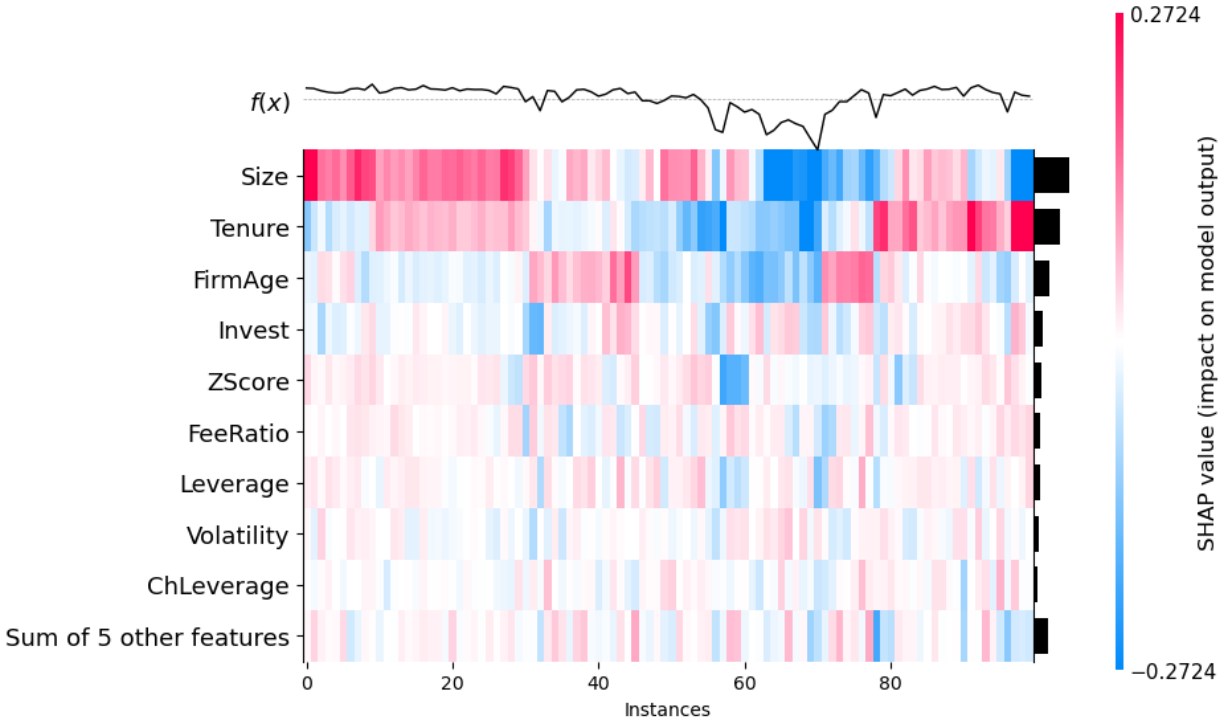
In [57]:
```
shap.plots.bar(shap_values.cohorts(2).abs.mean(0)) # Cohort plot
```



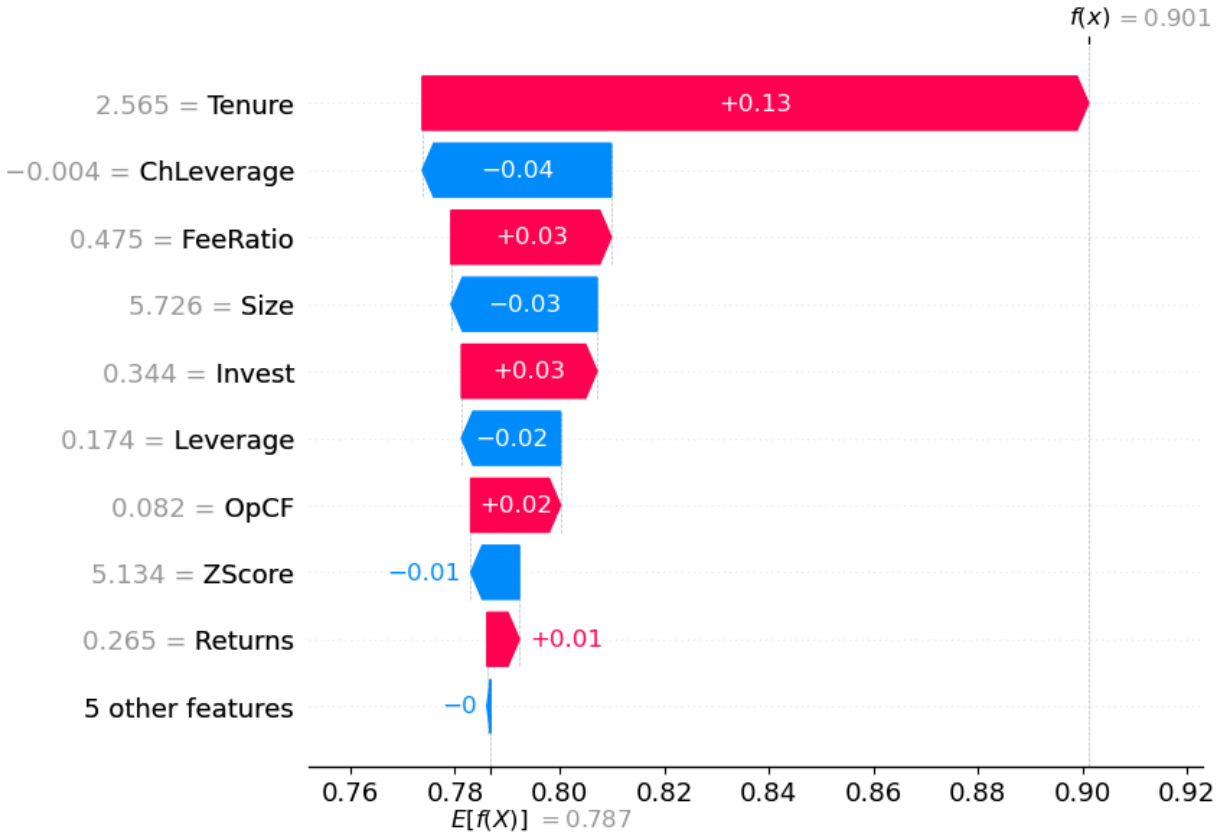
In [58]:
```
fig = plt.gcf() # gcf means "get current figure"
fig.set_figheight(11)
fig.set_figwidth(11)
shap.plots.heatmap(shap_values[1:100])
```



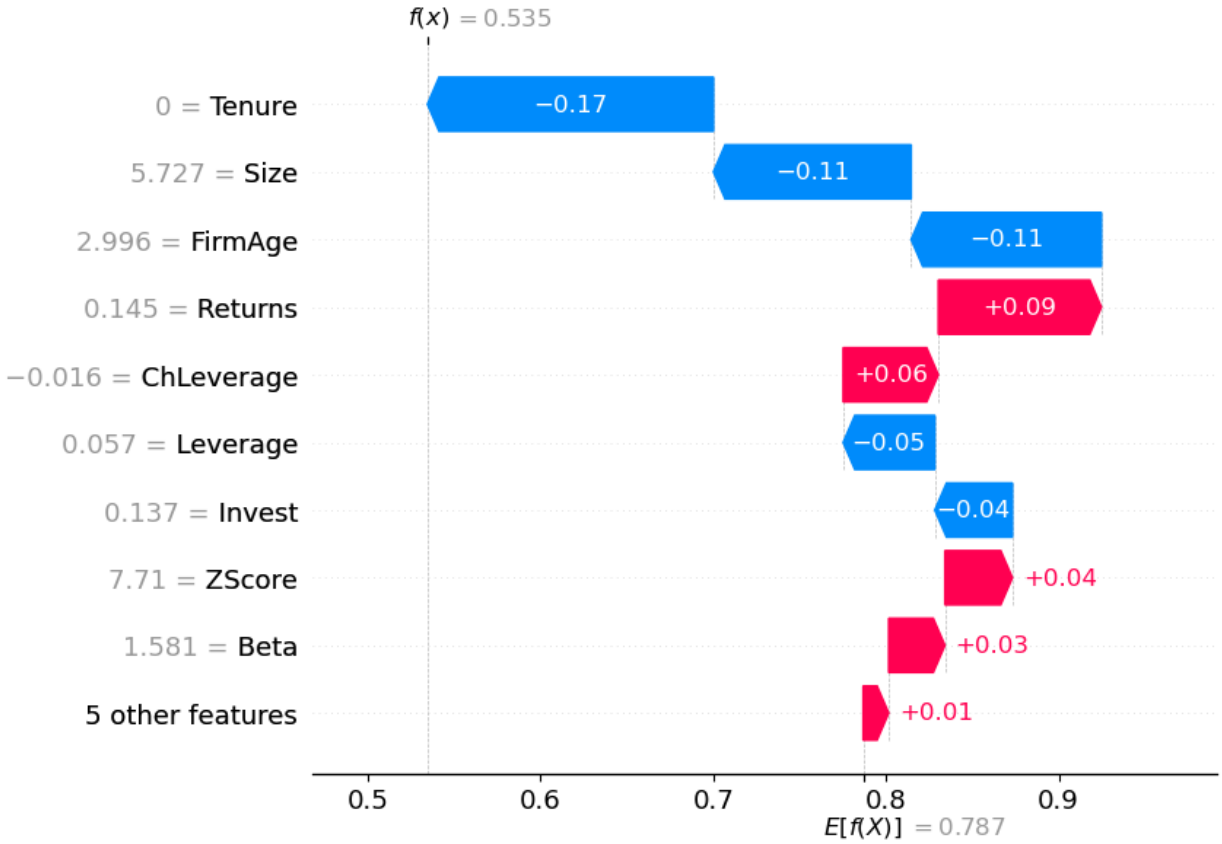
In [59]:
```
shap.plots.heatmap(shap_values[200:300])
```



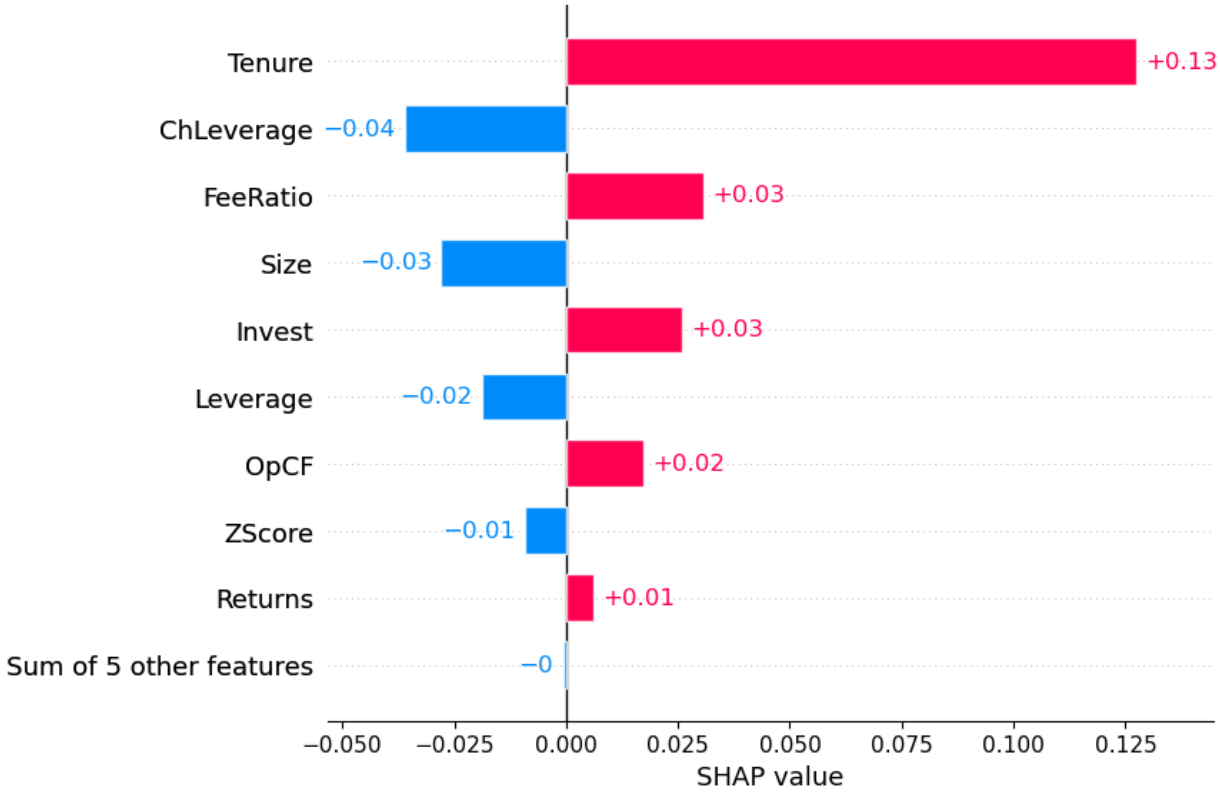
In [60]:
```
shap.plots.waterfall(shap_values[0]) # For the first observation
```



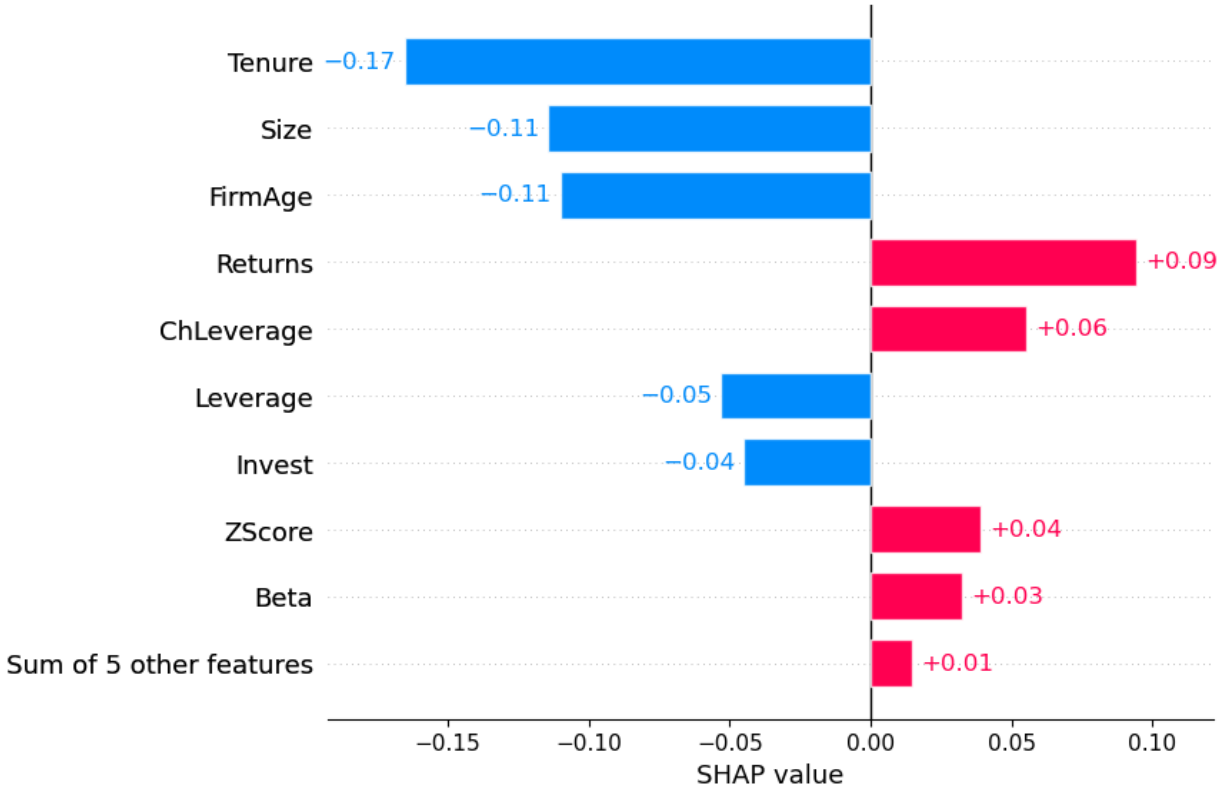
In [61]:
```
shap.plots.waterfall(shap_values[1]) # For the second observation
```



In [62]:
```
shap.plots.bar(shap_values[0]) # For the first observation barplot
```

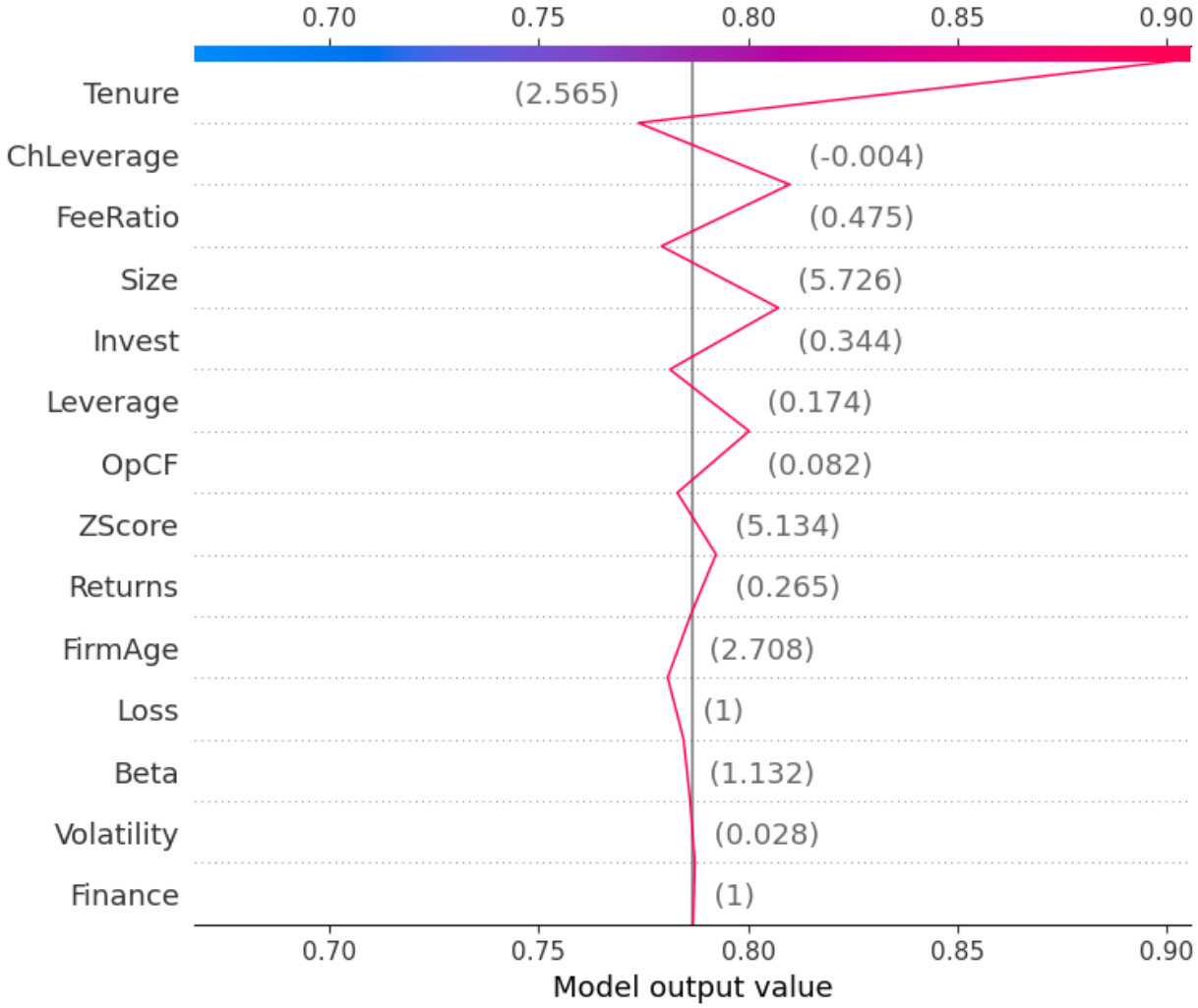


In [63]:
```
shap.plots.bar(shap_values[1]) # For the second observation barplot
```



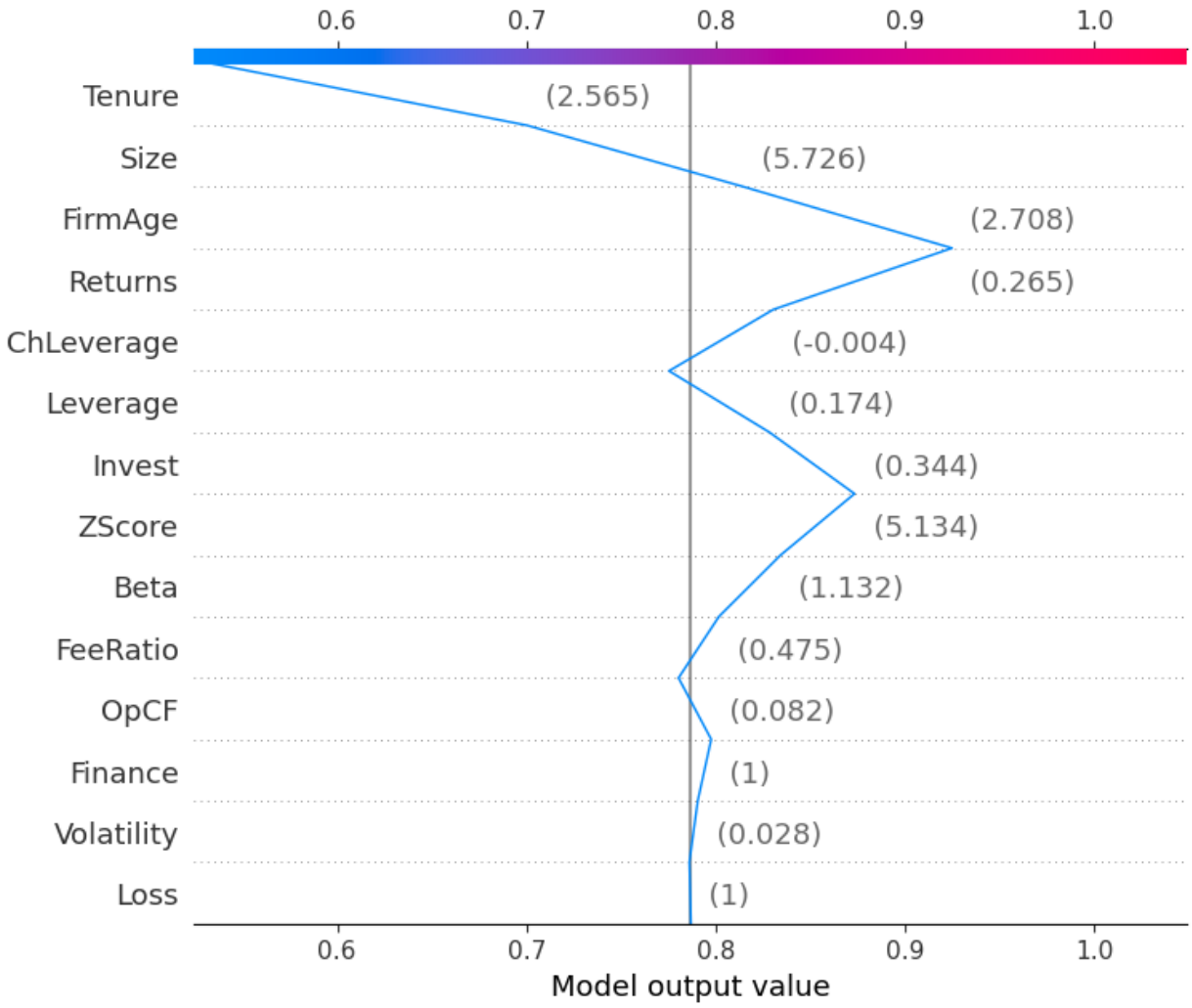
In [64]:
```
# Decision plot
expected_value = explainer.expected_value
#print("The expected value is ", expected_value)
print("The final prediction is ", xgb_model.predict(X_test)[0])
shap_values = explainer.shap_values(X_test)[0]
shap.decision_plot(expected_value, shap_values, X_test)
```

```
The final prediction is  0.9011796
```



In [65]:
```
shap_values = explainer.shap_values(X_test)[1]
#print("The expected value is ", expected_value)
print("The final prediction is ", xgb_model.predict(X_test)[1])
shap.decision_plot(expected_value, shap_values, X_test)
```

```
The final prediction is  0.5346452
```



In [66]:
```
# Binary Target
xgb_binary_model = xgb.XGBRegressor(objective='reg:logistic',random_state=42)
xgb_binary_model.fit(X_train, Y_train)
Y_pred = xgb_binary_model.predict(X_train)[0]
Y_pred
```

Out[66]:
```
0.99997056
```

In [67]:
```
# Waterfall plot
explainer = shap.Explainer(xgb_binary_model)
xgb_binary_shap_values = explainer(X_train)
```

```
[17:27:00] WARNING: /Users/runner/work/xgboost/xgboost/src/c_api/c_api.cc:124
0: Saving into deprecated binary model format, please consider using `json` or
`ubj`. Model format will default to JSON in XGBoost 2.2 if not specified.
[17:27:01] WARNING: /Users/runner/work/xgboost/xgboost/src/c_api/c_api.cc:124
0: Saving into deprecated binary model format, please consider using `json` or
`ubj`. Model format will default to JSON in XGBoost 2.2 if not specified.
```

In [68]:
```
from scipy.special import expit
expit(9.676)
```

Out[68]:
```
0.9999372318494586
```

In [69]:
```
original_shap_values = xgb_binary_shap_values
original_shap_values
```

Out[69]:
```
.values =
array([[ 1.5777497 ,  0.34928417, -0.38742083, ...,  0.35079858,
         0.06024942, -0.14367965],
       [ 2.7337563 ,  1.8042798 , -0.5370287 , ...,  0.1392942 ,
         0.13485812,  0.10250279],
       [-1.4035889 ,  0.03870981, -1.2086381 , ..., -0.15479349,
        -0.35088238, -0.23727198],
       ...,
       [-0.5556149 , -0.2396371 , -1.6877143 , ..., -0.16619559,
        -0.14312561,  0.05530791],
       [ 1.8892435 ,  0.5555458 , -0.41296577, ..., -0.2661265 ,
         0.98607045, -0.00834135],
       [-1.0717726 , -0.41847277,  1.1942096 , ..., -0.26563877,
         0.5042884 ,  0.8585152 ]], dtype=float32)

.base_values =
array([1.5671108, 1.5671108, 1.5671108, ..., 1.5671108, 1.5671108,
       1.5671108], dtype=float32)

.data =
array([[2.3025851 , 0.05320805, 2.99573231, ..., 0.27531603, 0.96938528,
        0.01808455],
       [2.8903718 , 0.36467293, 3.17805386, ..., 0.45569211, 1.44153595,
        0.01213791],
       [0.6931472 , 0.22529149, 3.13549423, ..., 0.40078792, 0.78609472,
        0.01260973],
       ...,
       [1.9459101 , 0.05312637, 2.94443893, ..., 0.41147998, 1.55957085,
        0.02640458],
       [2.5649493 , 0.07708553, 3.33220458, ..., 0.72826647, 2.23370677,
        0.03271395],
       [1.0986123 , 0.02189781, 1.38629436, ..., 0.7057573 , 2.23278385,
        0.02153001]])
```

In [70]:
```
# Compute the transformed base value, which consists in applying the logit fun
from scipy.special import expit # Importing the logit function for the base va
untransformed_base_value = original_shap_values.base_values[-1]
untransformed_base_value
```

Out[70]:
```
1.5671108
```

In [71]:
```
original_shap_values.values
```

Out[71]:
```
array([[ 1.5777497 ,  0.34928417, -0.38742083, ...,  0.35079858,
         0.06024942, -0.14367965],
       [ 2.7337563 ,  1.8042798 , -0.5370287 , ...,  0.1392942 ,
         0.13485812,  0.10250279],
       [-1.4035889 ,  0.03870981, -1.2086381 , ..., -0.15479349,
        -0.35088238, -0.23727198],
       ...,
       [-0.5556149 , -0.2396371 , -1.6877143 , ..., -0.16619559,
        -0.14312561,  0.05530791],
       [ 1.8892435 ,  0.5555458 , -0.41296577, ..., -0.2661265 ,
         0.98607045, -0.00834135],
       [-1.0717726 , -0.41847277,  1.1942096 , ..., -0.26563877,
         0.5042884 ,  0.8585152 ]], dtype=float32)
```

In [76]:
```
import numpy as np
original_explanation_distance = np.sum(original_shap_values.values, axis=1)[2]
```

```
original_explanation_distance
```

Out[76]:
```
-4.2045155
```

In [78]:
```
distance_to_explain = abs(Y_pred - untransformed_base_value)
distance_to_explain
```

Out[78]:
```
0.5671402
```

In [80]:
```
distance_coefficient = np.abs(original_explanation_distance / distance_to_expla
distance_coefficient
```

Out[80]:
```
7.413538
```

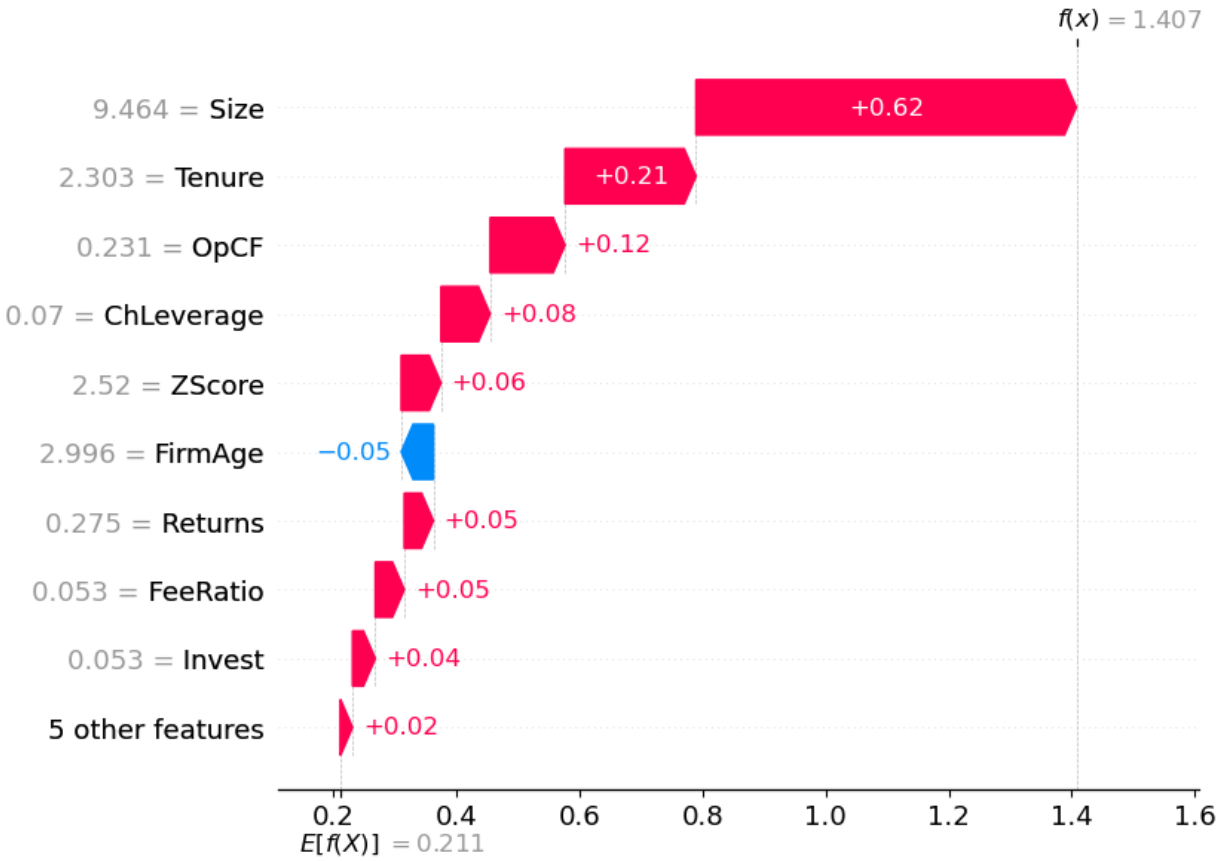
In [81]:
```
original_shap_values.values[2]
```

Out[81]:
```
array([-1.4035889e+00,  3.8709812e-02, -1.2086381e+00, -1.5047621e+00,
        1.7621313e+00, -4.8217103e-02, -6.5158802e-01,  1.2490553e-03,
       -2.8659341e-01, -2.4763167e-02, -1.3550682e-01, -1.5479349e-01,
       -3.5088238e-01, -2.3727198e-01], dtype=float32)
```

In [82]:
```
shap_values_transformed = original_shap_values / distance_coefficient
shap_values_transformed
```

Out[82]:
```
.values =
array([[ 0.21282008,  0.04711437, -0.05225856, ...,  0.04731864,
         0.00812695, -0.01938071],
       [ 0.3687519 ,  0.24337634, -0.07243892, ...,  0.01878917,
         0.01819079,  0.01382643],
       [-0.1893278 ,  0.0052215 , -0.1630312 , ..., -0.02087984,
        -0.04732995, -0.03200523],
       ...,
       [-0.07494599, -0.03232426, -0.22765303, ..., -0.02241785,
        -0.01930598,  0.00746039],
       [ 0.25483695,  0.07493667, -0.05570428, ..., -0.03589737,
         0.13300943, -0.00112515],
       [-0.14456965, -0.05644711,  0.16108498, ..., -0.03583158,
         0.06802263,  0.11580371]], dtype=float32)

.base_values =
array([0.21138501, 0.21138501, 0.21138501, ..., 0.21138501, 0.21138501,
       0.21138501], dtype=float32)

.data =
array([[0.31059193, 0.00717715, 0.40408943, ..., 0.03713693, 0.13075879,
        0.0024394 ],
       [0.38987752, 0.04919013, 0.42868248, ..., 0.06146756, 0.19444642,
        0.00163726],
       [0.09349749, 0.0303892 , 0.42294168, ..., 0.05406163, 0.10603503,
        0.00170091],
       ...,
       [0.26248063, 0.00716613, 0.39717055, ..., 0.05550386, 0.21036796,
        0.00356167],
       [0.34598181, 0.01039794, 0.44947562, ..., 0.09823467, 0.30130105,
        0.00441273],
       [0.14819001, 0.00295376, 0.18699498, ..., 0.09519845, 0.30117656,
        0.00290415]])
```

In [83]:
```
base_value = shap_values_transformed.base_values
shap_values_transformed.data = original_shap_values.data
shap.plots.waterfall(shap_values_transformed[0])
```



In [85]:
```
def xgb_shap_transform_scale(original_shap_values, Y_pred, which):
 from scipy.special import expit

 # Compute the transformed base value, which consists in applying the logit
 from scipy.special import expit #Importing the logit function for the base
 untransformed_base_value = original_shap_values.base_values[-1]

 # Computing the original_explanation_distance to construct the distance_coe
 original_explanation_distance = np.sum(original_shap_values.values, axis=1)

 base_value = expit(untransformed_base_value ) # = 1 / (1+ np.exp(-untransfo
 # Computing the distance between the model_prediction and the transformed b
 distance_to_explain = Y_pred[which] - base_value
 # The distance_coefficient is the ratio between both distances which will b
 distance_coefficient = original_explanation_distance / distance_to_explain
 # Transforming the original shapley values to the new scale
 shap_values_transformed = original_shap_values / distance_coefficient
 # Finally resetting the base_value as it does not need to be transformed
 shap_values_transformed.base_values = base_value
 shap_values_transformed.data = original_shap_values.data

 # Now returning the transformed array
 return shap_values_transformed
```

In [90]:
```
obs = 0
Y_pred = xgb_binary_model.predict(X_train)
print("The prediction is ", Y_pred[obs])
```

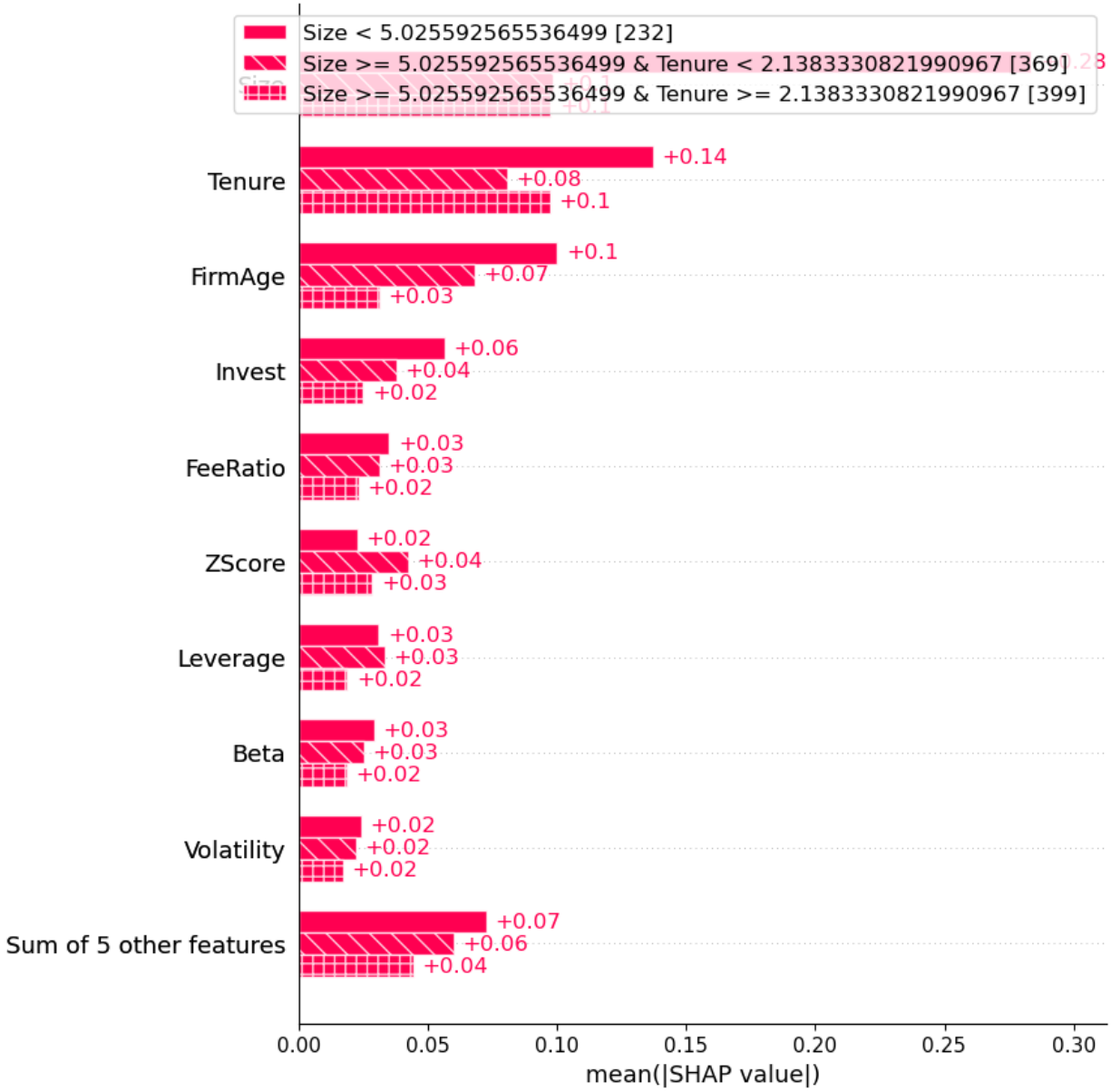
```
shap_values_transformed = xgb_shap_transform_scale(xgb_binary_shap_values, Y_p
shap.plots.waterfall(shap_values_transformed[obs])
```

```
The prediction is  0.99997056
```

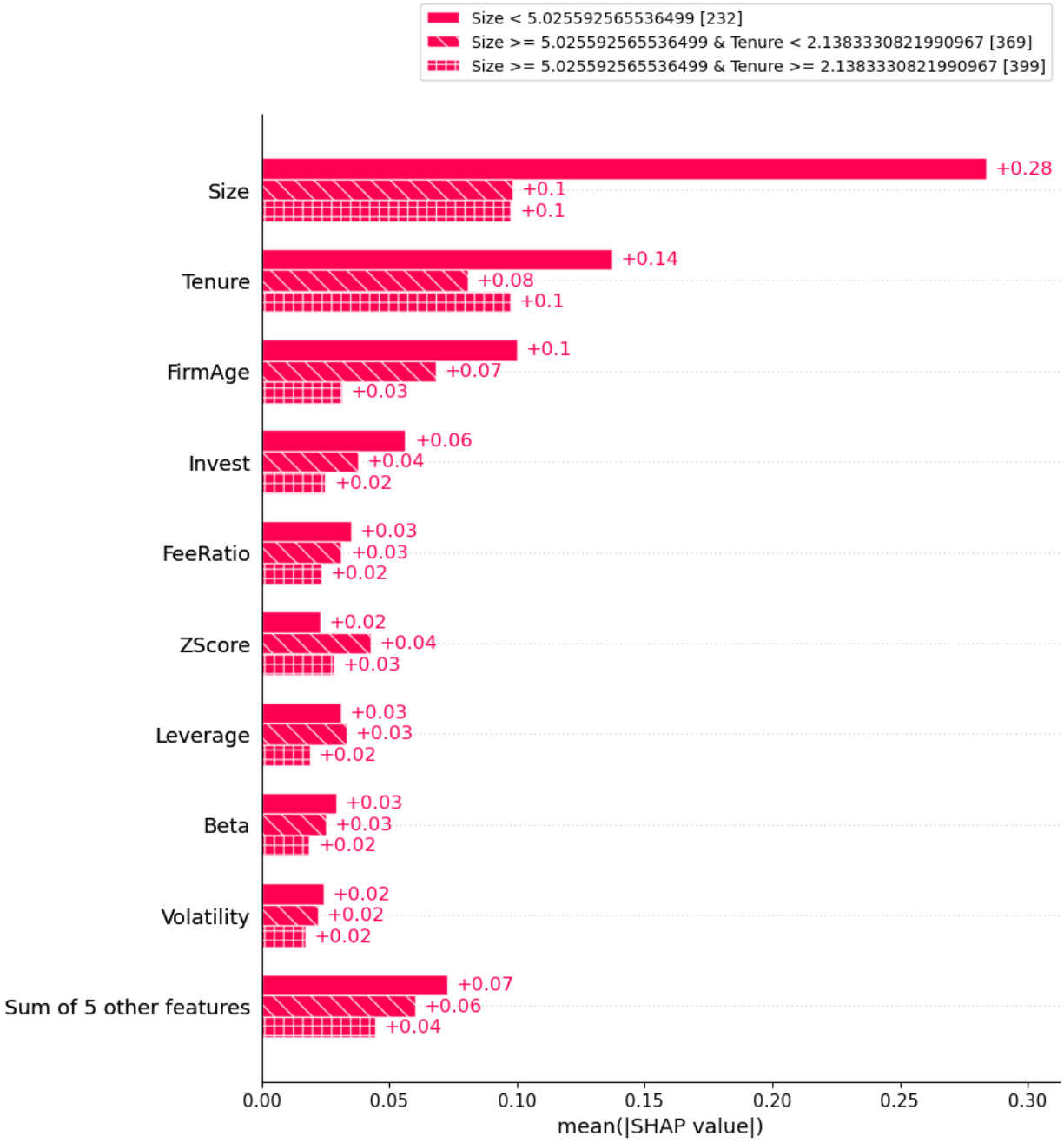
```
---------------------------------------------------------------------------
ValueError                                Traceback (most recent call last)
Cell In[90], line 4
      2 Y_pred = xgb_binary_model.predict(X_train)
      3 print("The prediction is ", Y_pred[obs])
----> 4 shap_values_transformed = xgb_shap_transform_scale(xgb_binary_shap_val
ues, Y_pred, 0)
      5 shap.plots.waterfall(shap_values_transformed[obs])

Cell In[85], line 17, in xgb_shap_transform_scale(original_shap_values, Y_pre
d, which)
     15 distance_coefficient = original_explanation_distance / distance_to_exp
lain
     16 # Transforming the original shapley values to the new scale
---> 17 shap_values_transformed = original_shap_values / distance_coefficient
     18 # Finally resetting the base_value as it does not need to be transform
ed
     19 shap_values_transformed.base_values = base_value

File ~/anaconda3/lib/python3.10/site-packages/shap/_explanation.py:495, in Exp
lanation.__truediv__(self, other)
    494 def __truediv__(self, other):
--> 495     return self._apply_binary_operator(other, operator.truediv, "__tru
ediv__")

File ~/anaconda3/lib/python3.10/site-packages/shap/_explanation.py:469, in Exp
lanation._apply_binary_operator(self, other, binary_op, op_name)
    467         new_exp.base_values = binary_op(new_exp.base_values, other.bas
e_values)
    468 else:
--> 469     new_exp.values = binary_op(new_exp.values, other)
    470     if new_exp.data is not None:
    471         new_exp.data = binary_op(new_exp.data, other)

ValueError: operands could not be broadcast together with shapes (4000,14) (40
00,)
```

In [91]:
```
explainer = shap.Explainer(xgb_model)
shap_values = explainer(X_test)
shap.plots.bar(shap_values.cohorts(3).abs.mean(0))
```

```
[16:21:22] WARNING: /Users/runner/work/xgboost/xgboost/src/c_api/c_api.cc:124
0: Saving into deprecated binary model format, please consider using `json` or
`ubj`. Model format will default to JSON in XGBoost 2.2 if not specified.
```



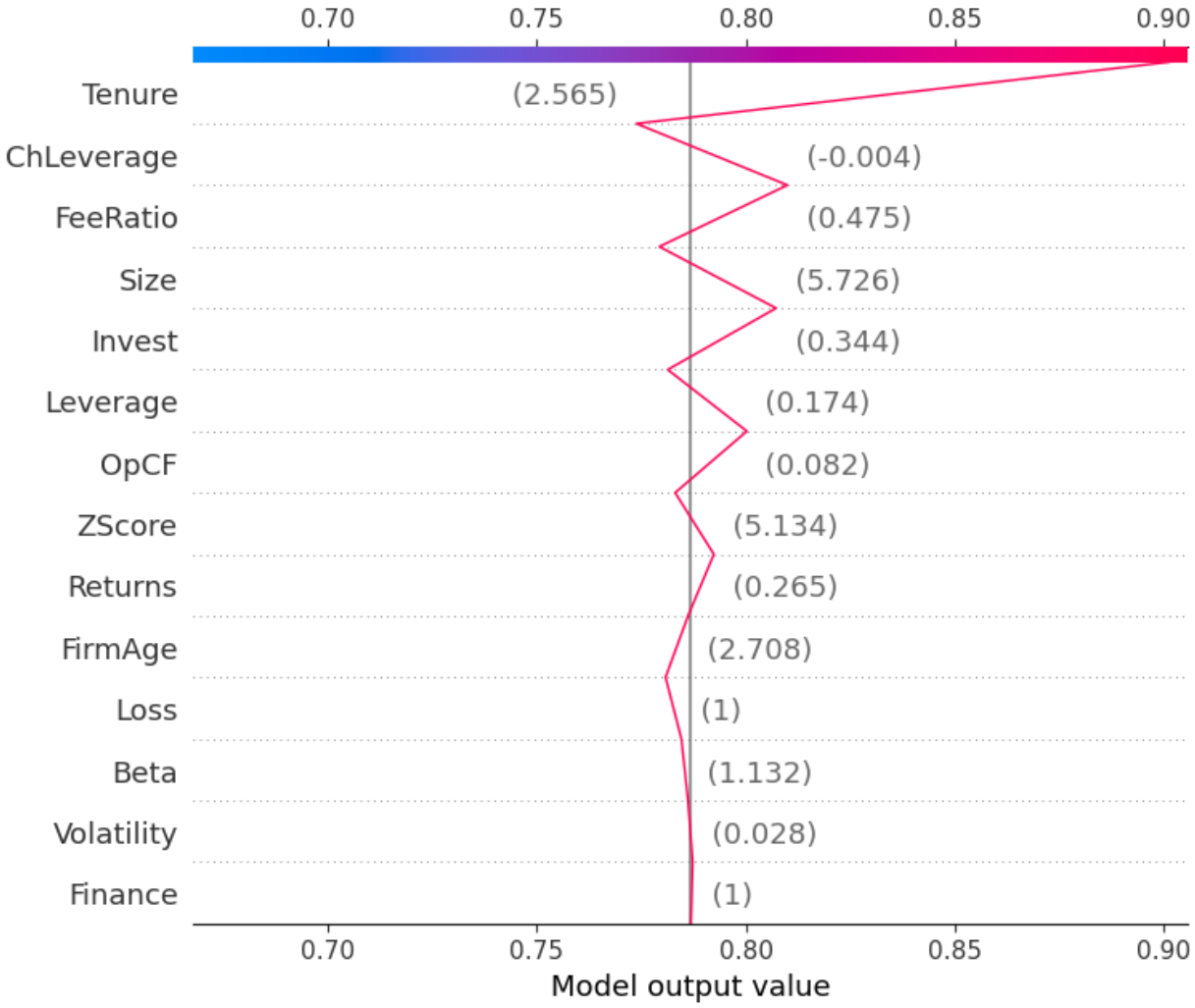
In [92]:
```
shap.plots.bar(shap_values.cohorts(3).abs.mean(0),show=False)
fig = plt.gcf() # gcf means "get current figure"
fig.set_figheight(11)
fig.set_figwidth(9)
#plt.rcParams['font.size'] = '12'
ax = plt.gca() #gca means "get current axes"
leg = ax.legend(bbox_to_anchor=(0., 1.02, 1., .102))
for l in leg.get_texts(): l.set_text(l.get_text().replace('Class', 'Klasse'))
plt.show()
```



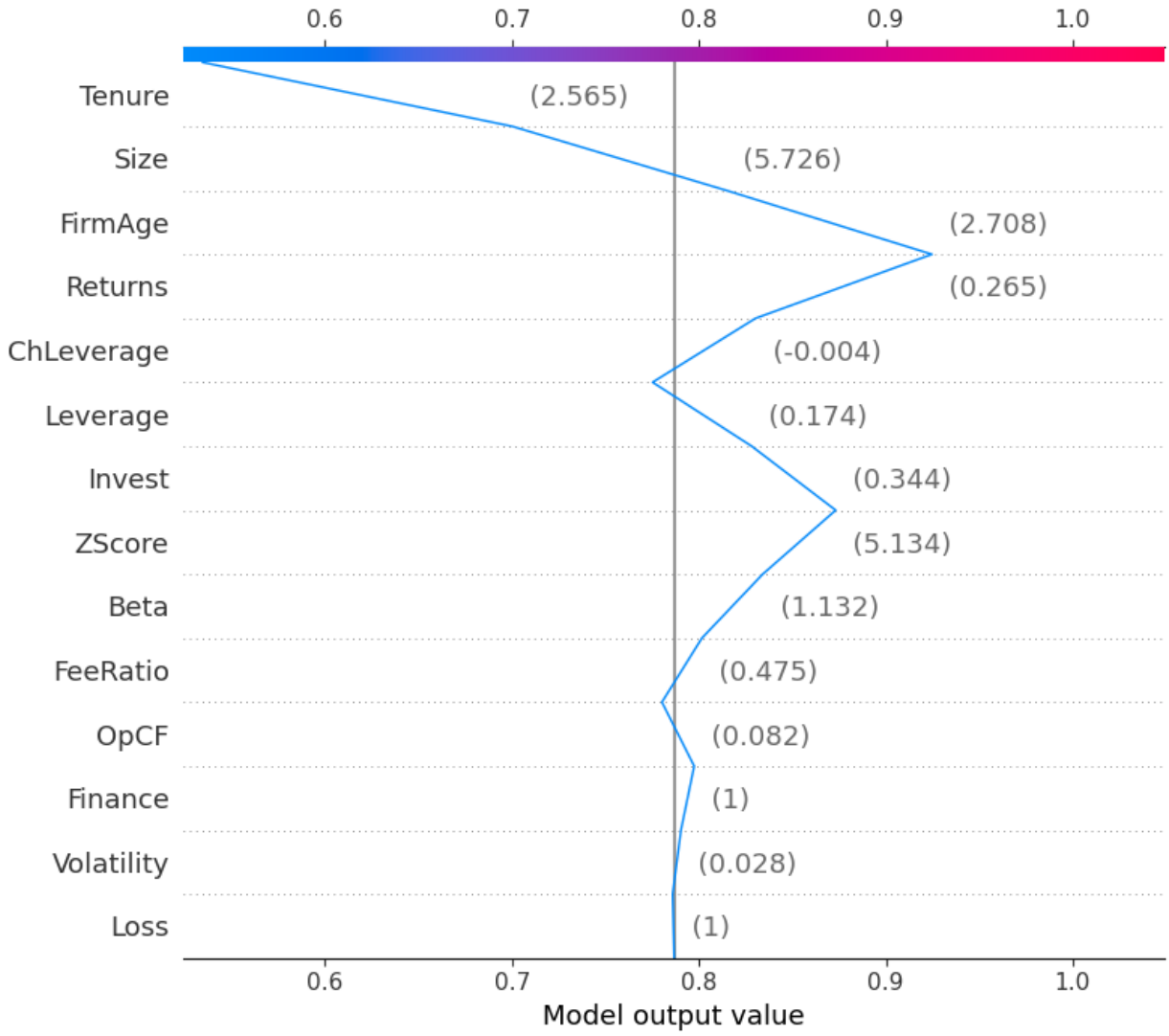
In [93]:
```
fig = plt.figure(figsize=(10,5))
```

```
<Figure size 1000x500 with 0 Axes>
```

In [94]:
```
ax1 = fig.add_subplot(121)
shap_values = explainer.shap_values(X_test)[0]
shap.decision_plot(expected_value, shap_values, X_test, show=False)
ax1.title.set_text('The First Observation')
```



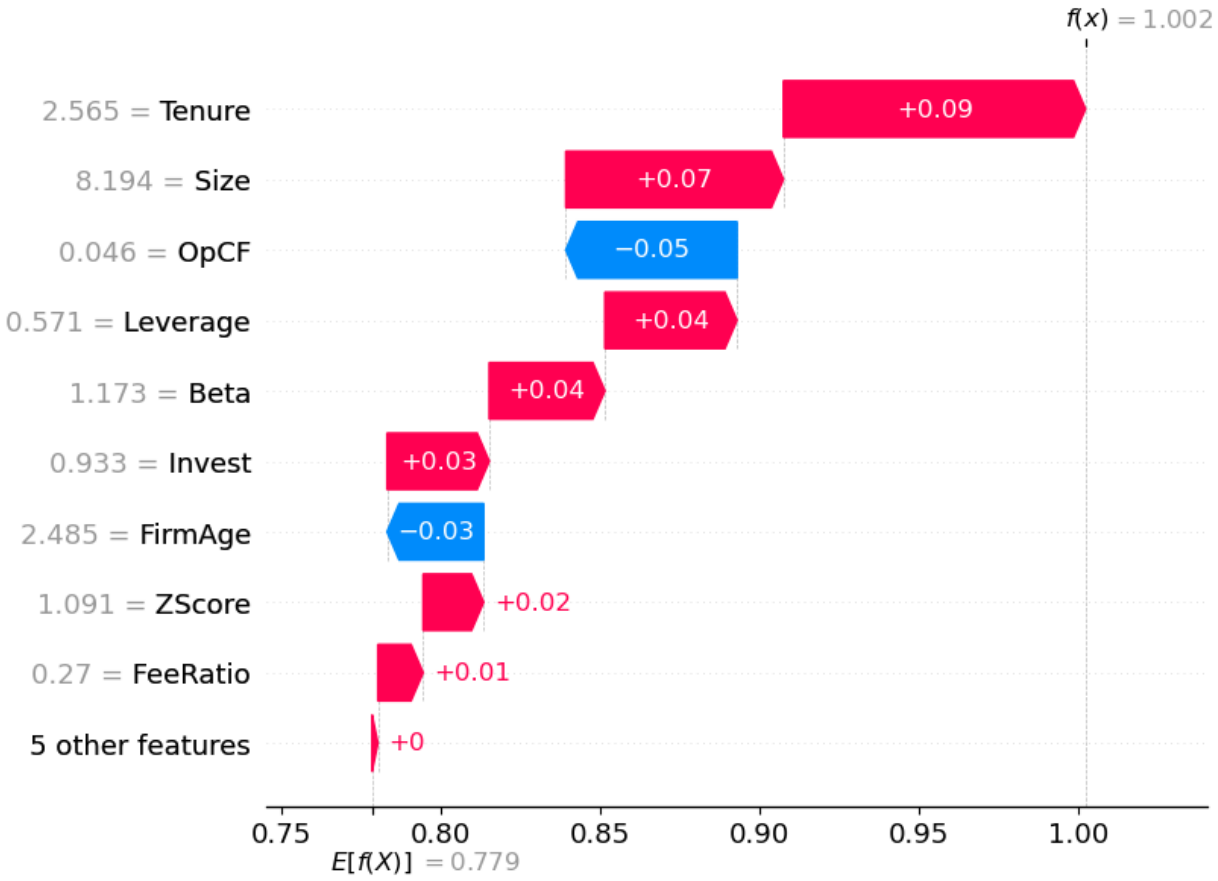
In [95]:
```
ax2 = fig.add_subplot(122)
shap_values = explainer.shap_values(X_test)[1]
shap.decision_plot(expected_value, shap_values, X_test, show=False)
ax2.title.set_text('The Second Observation')
plt.tight_layout()
plt.show()
```



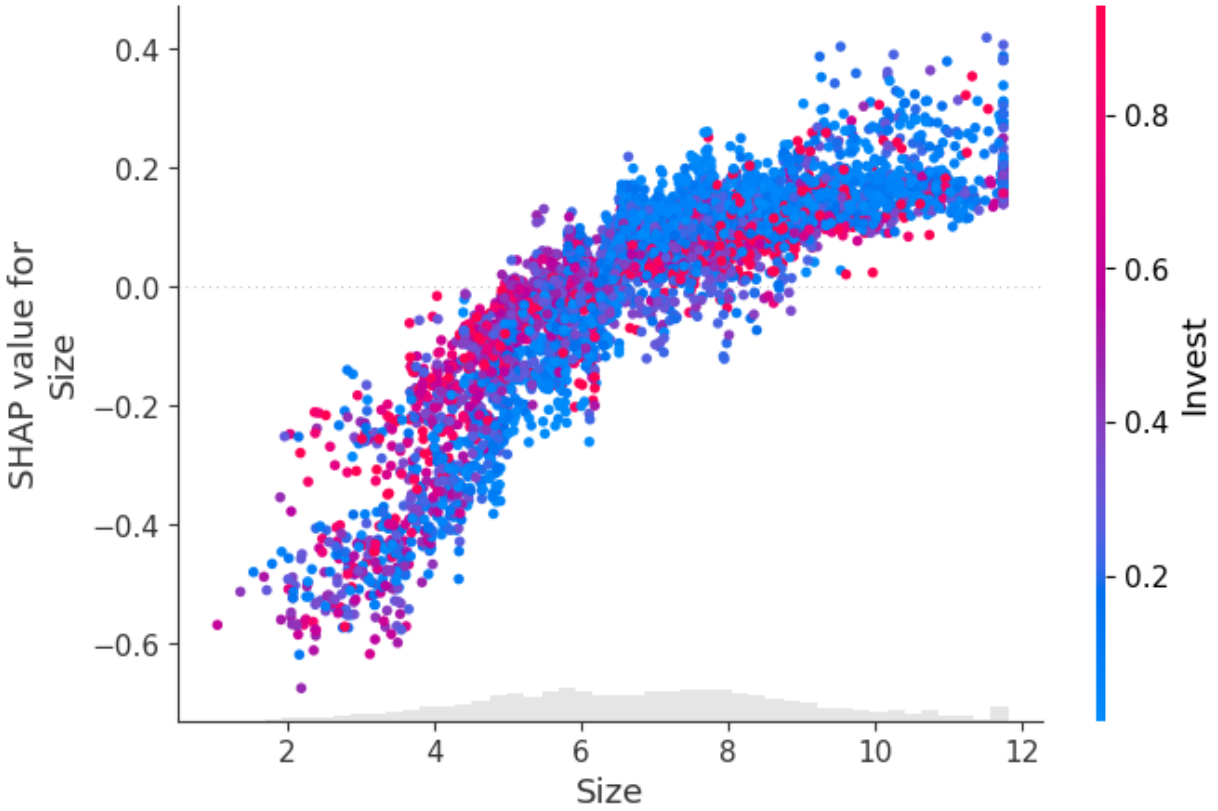
In [96]:
```
# train an XGBoost model
X, y = df[xvars],df['Big4']
model = xgb.XGBRegressor().fit(X, y)
# explain the model's predictions using SHAP
# (same syntax works for LightGBM, CatBoost, scikit-learn, transformers, Spark,
explainer = shap.Explainer(model)
shap_values = explainer(X)
```

```
[16:23:13] WARNING: /Users/runner/work/xgboost/xgboost/src/c_api/c_api.cc:124
0: Saving into deprecated binary model format, please consider using `json` or
`ubj`. Model format will default to JSON in XGBoost 2.2 if not specified.
```

In [97]:
```
# visualize the first prediction's explanation
shap.plots.waterfall(shap_values[0])
```

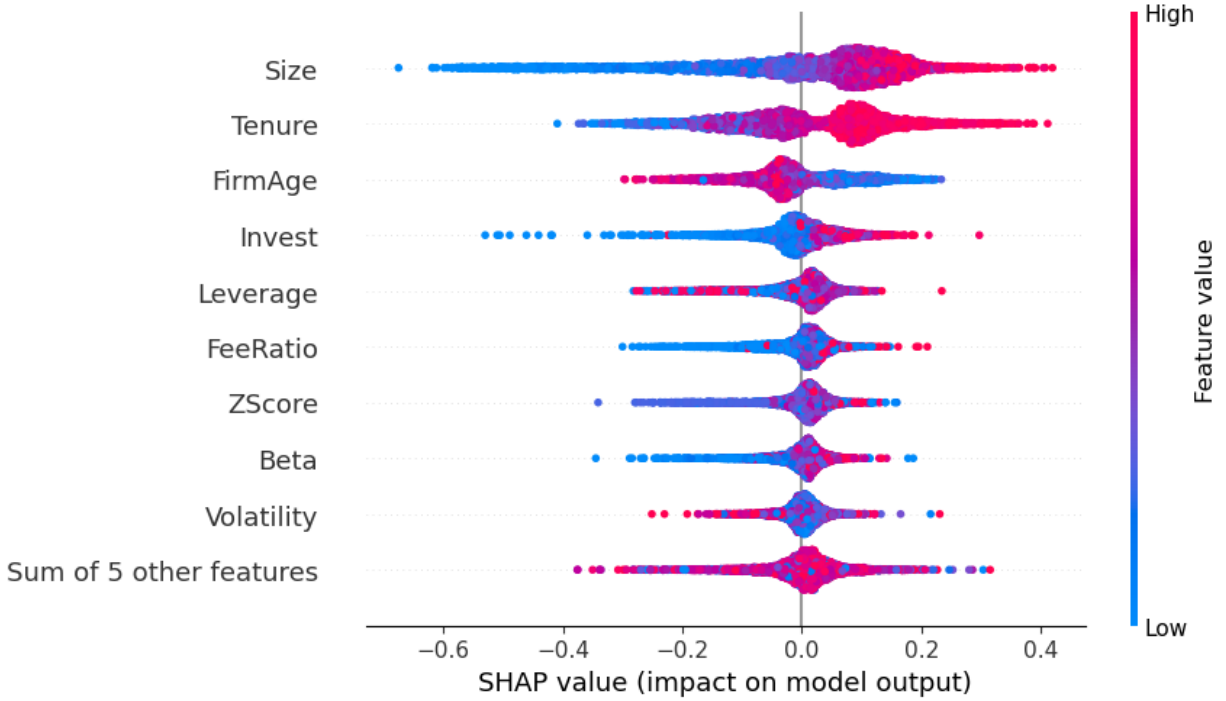


In [98]:
```
shap.plots.scatter(shap_values[:,"Size"], color=shap_values)
```



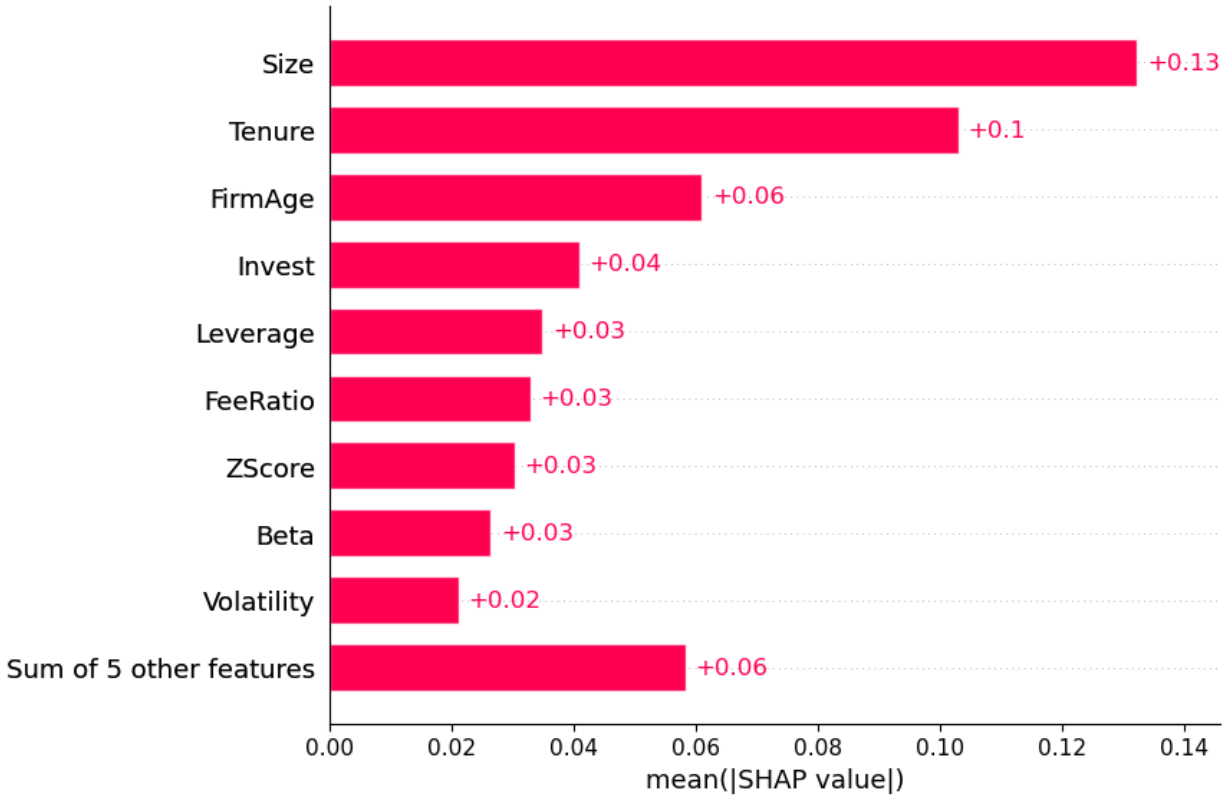
In [99]:
```
shap.plots.beeswarm(shap_values)
```

```
No data for colormapping provided via 'c'. Parameters 'vmin', 'vmax' will be i
gnored
```



In [100…

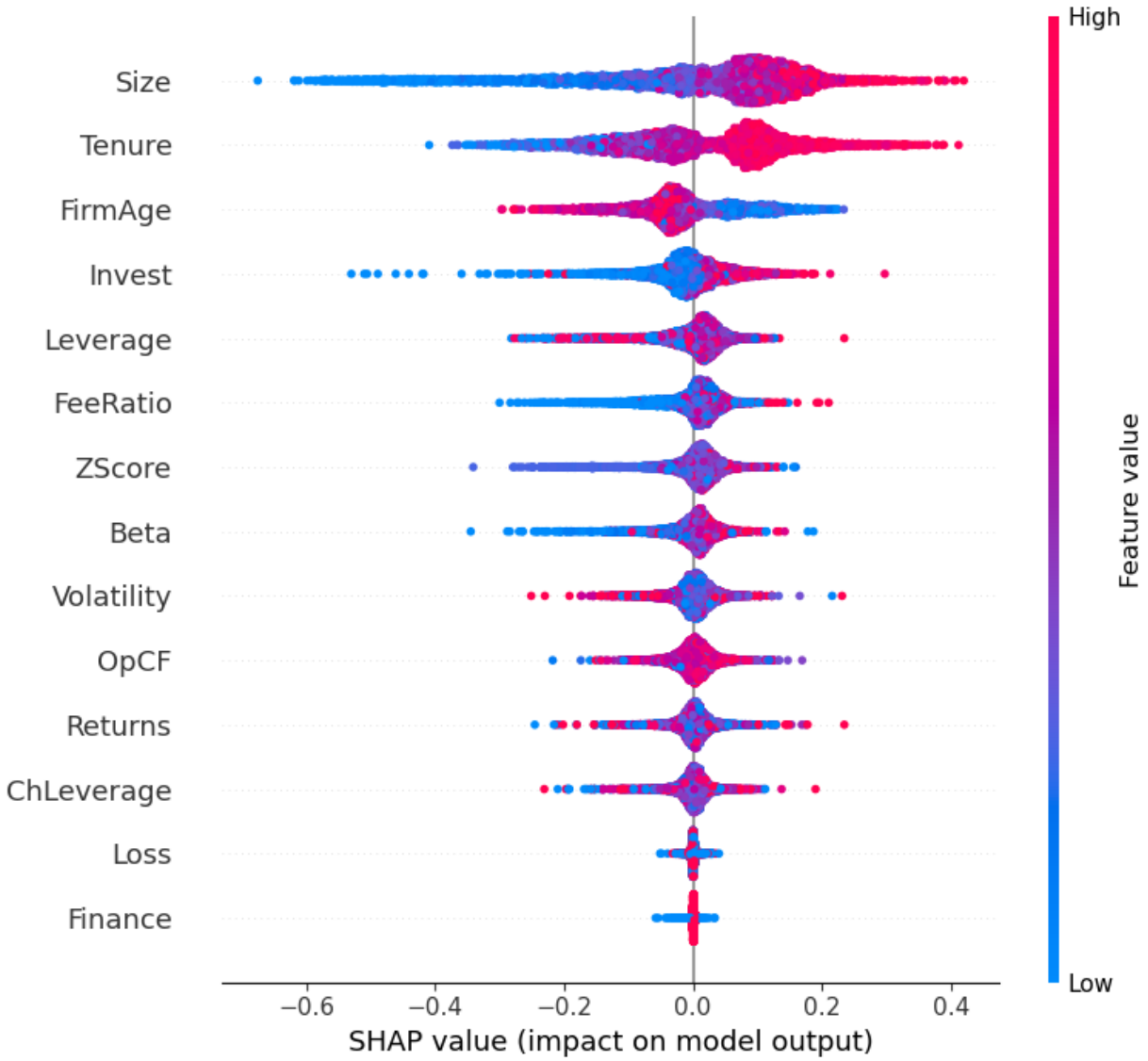
```
shap.plots.bar(shap_values)
```



In [101…

```
shap.summary_plot(shap_values, df[xvars])
```

```
No data for colormapping provided via 'c'. Parameters 'vmin', 'vmax' will be i
gnored
```



In [102…

```
print(df.describe().transpose())
df.describe().T
```

	count	mean	std	min	25%	50%	75%	max
GCO	5000.0	0.010600	0.102419	0.000000e+00	0.000000	0.000000	0.000000	1.000000
Tenure	5000.0	1.904949	0.751237	0.000000e+00	1.386294	2.079442	2.484907	2.890372
Big4	5000.0	0.778600	0.415231	0.000000e+00	1.000000	1.000000	1.000000	1.000000
FeeRatio	5000.0	0.141429	0.138196	0.000000e+00	0.024196	0.104314	0.222111	0.621101
FirmAge	5000.0	2.730144	0.752482	0.000000e+00	2.197225	2.772589	3.218876	4.174387
ZScore	5000.0	2.232858	6.745728	-1.471205e+02	0.484653	1.544287	3.151230	100.537613
Size	5000.0	6.898541	2.108985	1.058137e+00	5.368144	6.890536	8.316711	11.736827
Loss	5000.0	0.483600	0.499781	0.000000e+00	0.000000	0.000000	1.000000	1.000000
Leverage	5000.0	0.584121	0.283002	2.178798e-02	0.384587	0.590995	0.762653	2.729931
ChLeverage	5000.0	0.016008	0.157074	-6.009547e+00	-0.023092	0.007548	0.049285	2.395992
OpCF	5000.0	0.017839	0.189054	-3.164750e+00	-0.002316	0.052809	0.097786	0.359912
Finance	5000.0	0.925400	0.262771	0.000000e+00	1.000000	1.000000	1.000000	1.000000
Invest	5000.0	0.320069	0.299344	1.940000e-07	0.074824	0.209023	0.512577	0.995753
Returns	5000.0	0.414332	0.195552	3.283552e-02	0.296476	0.394603	0.488119	1.461594
Beta	5000.0	1.136815	0.571122	-4.444910e-01	0.750049	1.108556	1.473767	2.815003
Volatility	5000.0	0.025848	0.015291	7.749172e-03	0.015060	0.022943	0.032470	0.137875

Out[102]:

	count	mean	std	min	25%	50%	75%	
GCO	5000.0	0.010600	0.102419	0.000000e+00	0.000000	0.000000	0.000000	1.0
Tenure	5000.0	1.904949	0.751237	0.000000e+00	1.386294	2.079442	2.484907	2.8
Big4	5000.0	0.778600	0.415231	0.000000e+00	1.000000	1.000000	1.000000	1.0
FeeRatio	5000.0	0.141429	0.138196	0.000000e+00	0.024196	0.104314	0.222111	0.
FirmAge	5000.0	2.730144	0.752482	0.000000e+00	2.197225	2.772589	3.218876	4.
ZScore	5000.0	2.232858	6.745728	-1.471205e+02	0.484653	1.544287	3.151230	100.5
Size	5000.0	6.898541	2.108985	1.058137e+00	5.368144	6.890536	8.316711	11.7
Loss	5000.0	0.483600	0.499781	0.000000e+00	0.000000	0.000000	1.000000	1.0
Leverage	5000.0	0.584121	0.283002	2.178798e-02	0.384587	0.590995	0.762653	2.7
ChLeverage	5000.0	0.016008	0.157074	-6.009547e+00	-0.023092	0.007548	0.049285	2.3
OpCF	5000.0	0.017839	0.189054	-3.164750e+00	-0.002316	0.052809	0.097786	0.3
Finance	5000.0	0.925400	0.262771	0.000000e+00	1.000000	1.000000	1.000000	1.0
Invest	5000.0	0.320069	0.299344	1.940000e-07	0.074824	0.209023	0.512577	0.9
Returns	5000.0	0.414332	0.195552	3.283552e-02	0.296476	0.394603	0.488119	1.4
Beta	5000.0	1.136815	0.571122	-4.444910e-01	0.750049	1.108556	1.473767	2.8
Volatility	5000.0	0.025848	0.015291	7.749172e-03	0.015060	0.022943	0.032470	0.

In [105…

```
X, y = df[xvars],df['GCO']
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.25, ra
from sklearn.preprocessing import StandardScaler
sc=StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.fit_transform(X_test)
from sklearn.linear_model import LogisticRegression
classifier = LogisticRegression(random_state=0)
classifier.fit(X_train, y_train)
# print(classifier.predict(sc.transform([[30,87000]])))
y_pred=classifier.predict(X_test)
print(np.concatenate((y_pred,y_test)))
```

```
[0 0 0 ... 0 0 0]
```

In [106…

```
model = LogisticRegression(solver='liblinear', C=10, random_state=0).fit(X, y)
model.fit(X, y)
```

Out[106]:

```
LogisticRegression(C=10, random_state=0, solver='liblinear')
```

In [107…

```
model.intercept_
```

Out[107]:

```
array([-0.06138551])
```

In [108…

```
model.coef_
```

Out[108]:
```
array([[ 1.15792902e-01, -3.93672366e-01, -6.56029395e-01,
        -4.24949993e-03, -4.55376186e-01,  1.64317434e+00,
         7.42275404e-01,  4.60870433e-02, -1.59900583e+00,
        -4.63880680e-01, -9.23187666e-01, -4.41882120e+00,
        -4.47465708e-01,  3.67202122e+00]])
```

In [109…
```
model.predict_proba(X)
```

Out[109]:
```
array([[9.99874050e-01, 1.25950001e-04],
       [9.99424928e-01, 5.75071974e-04],
       [9.91115991e-01, 8.88400890e-03],
       ...,
       [9.98976165e-01, 1.02383506e-03],
       [9.99815810e-01, 1.84190342e-04],
       [9.99889245e-01, 1.10755283e-04]])
```

In [110…
```
model.score(X, y)
```

Out[110]:
```
0.9896
```

In [111…
```
import statsmodels.api as sm
model = sm.Logit(y, X)
result = model.fit(method='newton')
result.summary()
```

```
---------------------------------------------------------------------------
ModuleNotFoundError                       Traceback (most recent call last)
Cell In[111], line 1
----> 1 import statsmodels.api as sm
      2 model = sm.Logit(y, X)
      3 result = model.fit(method='newton')

ModuleNotFoundError: No module named 'statsmodels'
```

In [112…
```
X = sm.add_constant(X) # add a constant / intercept
model = sm.Logit(y, X)
result = model.fit(method='newton')
result.summary()
```

```
---------------------------------------------------------------------------
NameError                                 Traceback (most recent call last)
Cell In[112], line 1
----> 1 X = sm.add_constant(X) # add a constant / intercept
      2 model = sm.Logit(y, X)
      3 result = model.fit(method='newton')

NameError: name 'sm' is not defined
```

In [114…
```
from sklearn.tree import DecisionTreeClassifier
from sklearn.model_selection import train_test_split, RandomizedSearchCV
from sklearn.metrics import confusion_matrix, classification_report
params = {'criterion':['gini','entropy'],
 'max_depth':[10,50,100,None],
 'random_state':[123],
 'class_weight':[None,'balanced']}
dt = DecisionTreeClassifier() # Leave this as default
train,test = train_test_split(df,train_size=0.80, random_state=123)
# parameters: n_jobs = -1 - parallelize on all cores; cv=5 - 5-fold cross-val;
# n_iter = 10 - 10 different parameter sets; verbose = 5 - print
# some output
```

```
rs = RandomizedSearchCV(dt,param_distributions=params,n_jobs=-1,cv=5,n_iter=10
res = rs.fit(train[xvars],train['Big4'])
```

```
Fitting 5 folds for each of 10 candidates, totalling 50 fits
```

```
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
```

```
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
```

```
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
```

```
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
```

```
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
```

```
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
```

```
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
```

```
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
```

```
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
```

```
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
```

```
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
```

```
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
/Users/yeiserross/anaconda3/lib/python3.10/site-packages/sklearn/base.py:347:
InconsistentVersionWarning: Trying to unpickle estimator DecisionTreeClassifie
r from version 1.3.2 when using version 1.3.0. This might lead to breaking cod
e or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainabilit
y-limitations
  warnings.warn(
```

In [ ]:

```
print("Confusion matrix (Rows = True 0/1, Columns = Predicted 0/1):")
print(confusion_matrix(test['Big4'],rs.predict(test[xvars])))
print("\n\nscikit-learn's classification report, which gives information on acc
print(classification_report(test['Big4'],rs.predict(test[xvars])))
print("\n\nThe search identified this as the best parameters:")
print(res.best_params_)
```

In [ ]: