

Web3 Data Analytics

Sindhu Bommali

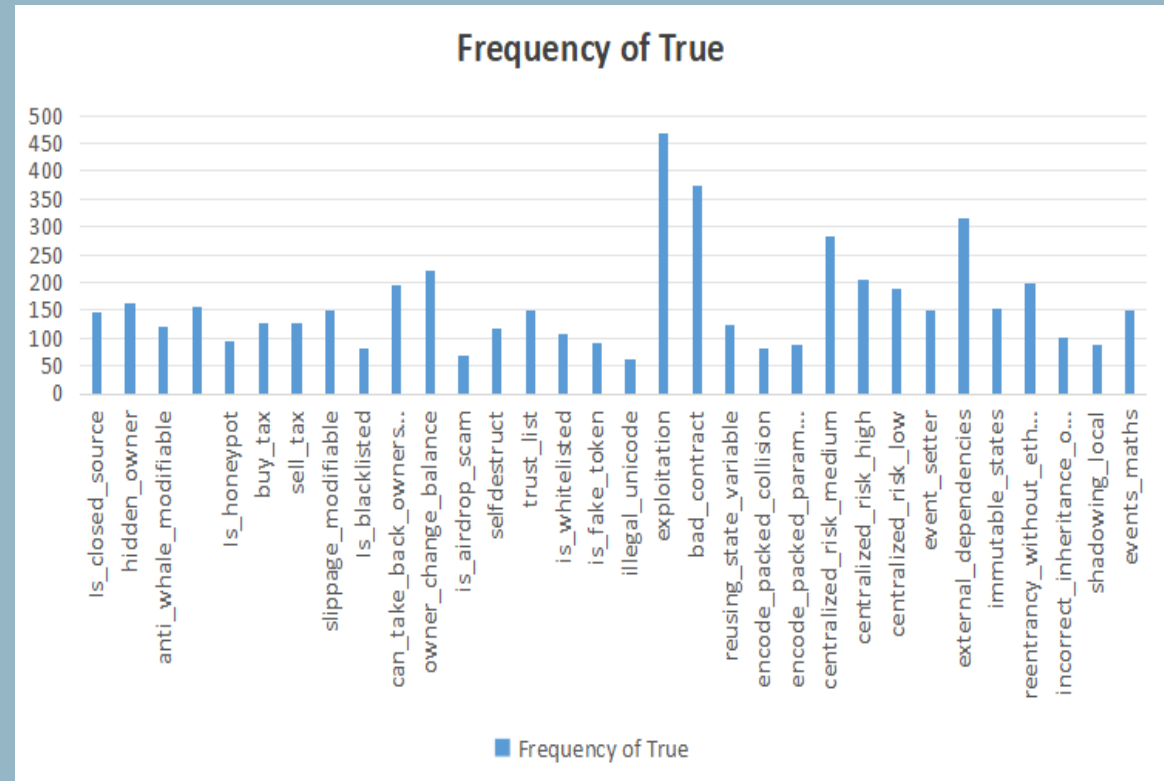
agenda

DATA PREPARATION & FREQUENCY
ANALYSIS

ADVANCED FREQUENCY ANALYSIS

CORRELATION ANALYSIS

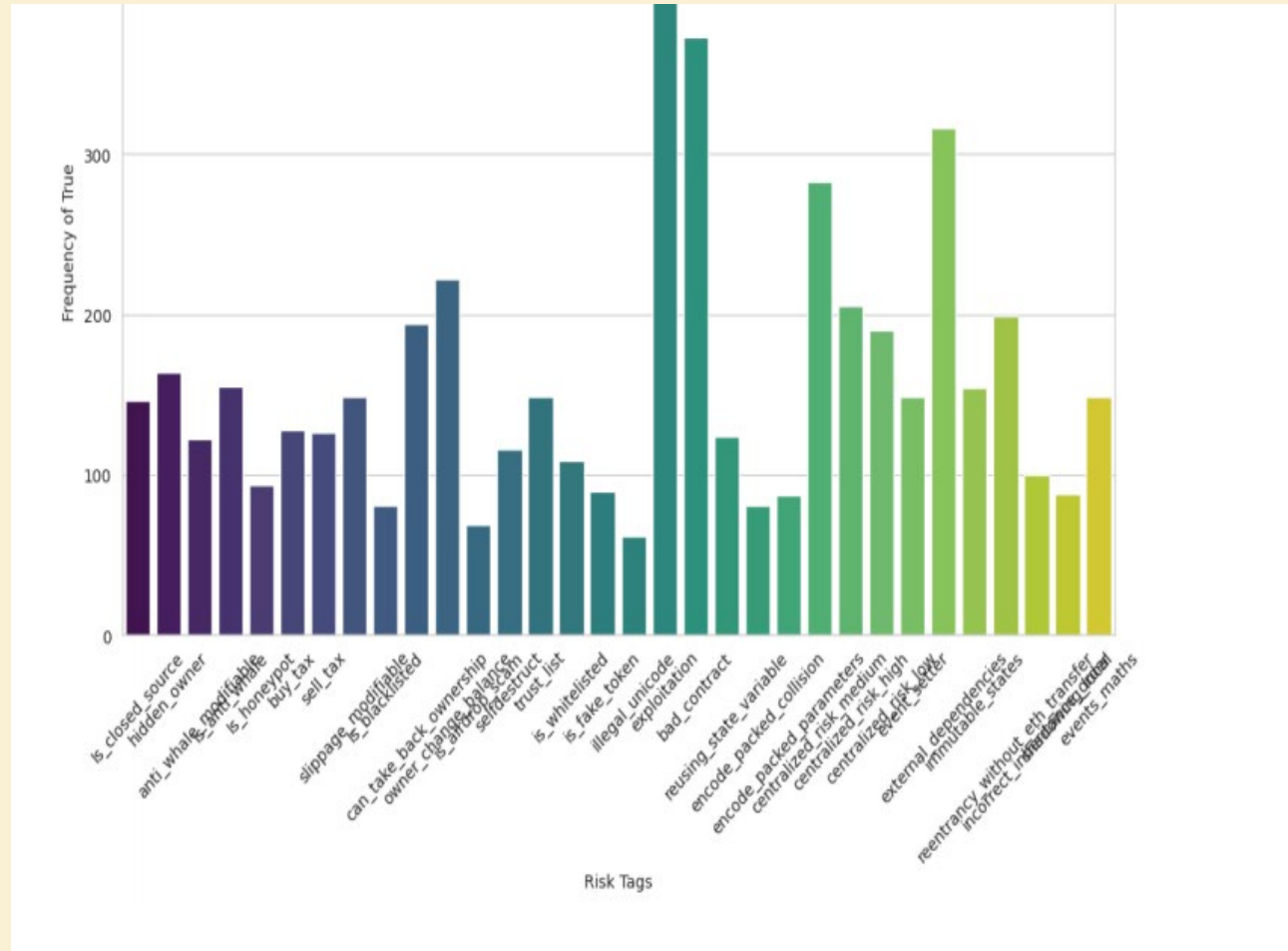
A detailed chart to represent Frequency Analysis of no. of risk tags marked true



Advanced Frequency Analysis of Smart Contracts

Is_closed_source	146
hidden_owner	164
anti_whale_modifiable	122
Is_anti_whale	155
Is_honeypot	94
buy_tax	128
sell_tax	126
slippage_modifiable	149
Is_blacklisted	81
can_take_back_ownership	194
owner_change_balance	222
is_airdrop_scam	69
selfdestruct	116
trust_list	149
is_whitelisted	109
is_fake_token	90
illegal_unicode	62
exploitation	468
bad_contract	373
reusing_state_variable	124
encode_packed_collision	81
encode_packed_parameters	87
centralized_risk_medium	283
centralized_risk_high	205
centralized_risk_low	190
event_setter	149
external_dependencies	316
immutable_states	154
reentrancy_without_eth_transfer	199
incorrect_inheritance_order	100
shadowing_local	88
events_maths	149

Name: True, dtype: int64



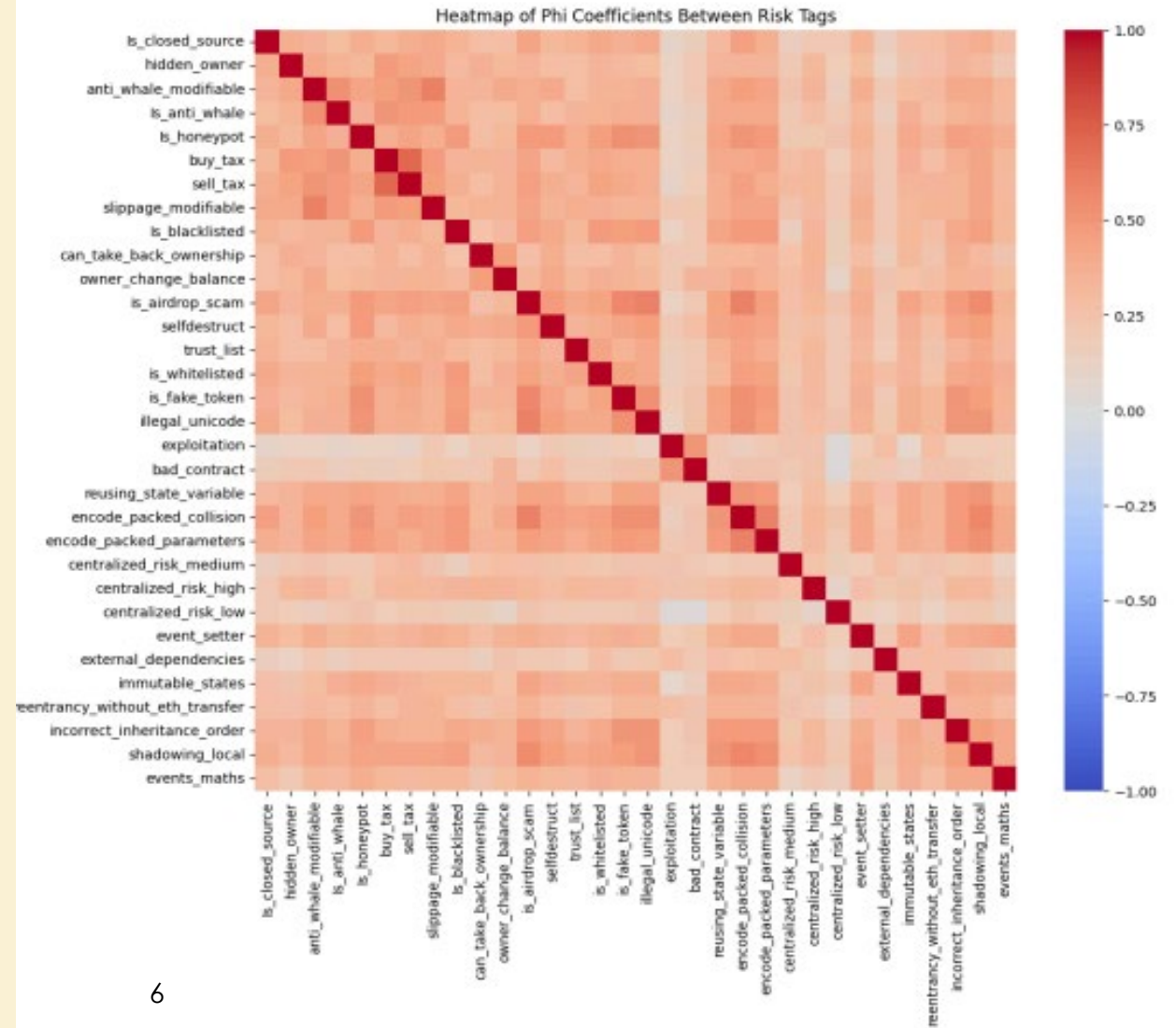
Correlation Analysis of Smart Contracts

```
def phi_coefficient(x, y):  
    """Calculate the Phi coefficient for two binary variables."""  
    # Create a contingency table  
    contingency_table = pd.crosstab(x, y)  
    # Calculate the phi coefficient from the contingency table  
    chi2 = scipy.stats.chi2_contingency(contingency_table, correction=False)[0]  
    n = np.sum(np.sum(contingency_table))  
    phi = np.sqrt(chi2 / n)  
    return phi  
  
# Example calculation between two risk tags  
phi = phi_coefficient(df['Is_honeypot'], df['anti_whale_modifiable'])  
print(f"Phi Coefficient between 'Is_honeypot' and 'anti_whale_modifiable': {phi}")  
  
→ Phi Coefficient between 'Is_honeypot' and 'anti_whale_modifiable': 0.43014356785902874
```

calculate the Phi coefficient, which is suitable for pairs of binary variables, we first need to establish a function that can handle this calculation

HeatMap Visualization of Smart Contract Risk tags

- Though we have the full correlation matrix in front of us, it is very difficult to visualize.
- Display those correlations where value is significantly positive or negative.



Thankyou

SINDHU BOMMALI