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

In [4]:
```
data =pd.read_csv("MI6_GW_MTW_GA.csv");
```

In [6]:
```
data
```

Out[6]:

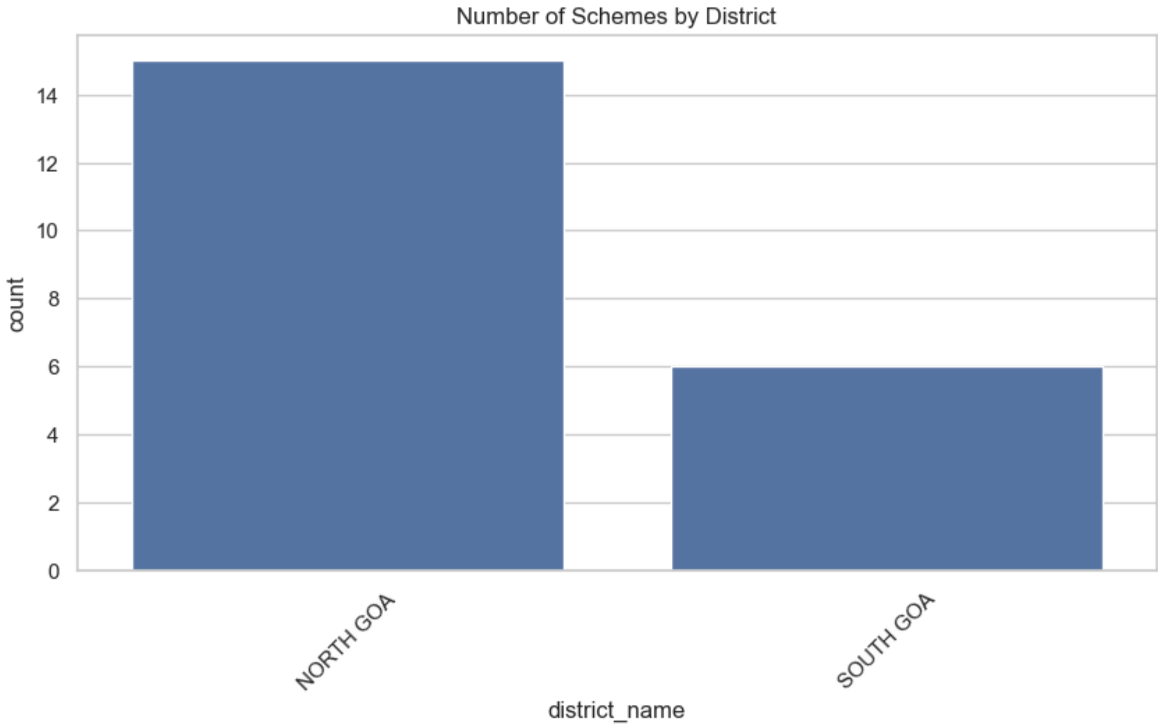
	state_name	district_name	block_tehsil_name	village_name	s_no_of_scheme	schem
0	GOA	NORTH GOA	SATARI	ADVOI	7	
1	GOA	NORTH GOA	SATARI	AMBEDEM	17	
2	GOA	NORTH GOA	SATARI	BIRONDEM	3	
3	GOA	NORTH GOA	SATARI	BIRONDEM	2	
4	GOA	NORTH GOA	SATARI	DONGURLI	11	
5	GOA	NORTH GOA	SATARI	PALE	8	
6	GOA	NORTH GOA	SATARI	PALE	13	
7	GOA	NORTH GOA	SATARI	PISSURLEM	4	
8	GOA	NORTH GOA	SATARI	AMBEDEM	13	
9	GOA	NORTH GOA	PERNEM	TORXEM	10	
10	GOA	NORTH GOA	PERNEM	TORXEM	17	
11	GOA	NORTH GOA	PERNEM	ARAMBOL	33	
12	GOA	SOUTH GOA	PONDA	QUERIM	83	
13	GOA	SOUTH GOA	DHARBANDORA	PILIEM	2	
14	GOA	SOUTH GOA	DHARBANDORA	PILIEM	3	
15	GOA	SOUTH GOA	QUEPEM	ASSOLDA	4	
16	GOA	NORTH GOA	SATARI	CODIEM	2	
17	GOA	SOUTH GOA	SANGUEM	RIVONA	19	
18	GOA	SOUTH GOA	DHARBANDORA	SIGAO	11	
19	GOA	NORTH GOA	SATARI	SIGONEM	3	

	state_name	district_name	block_tehsil_name	village_name	s_no_of_scheme	schem
20	GOA	NORTH GOA	SATARI	SIGONEM	4	

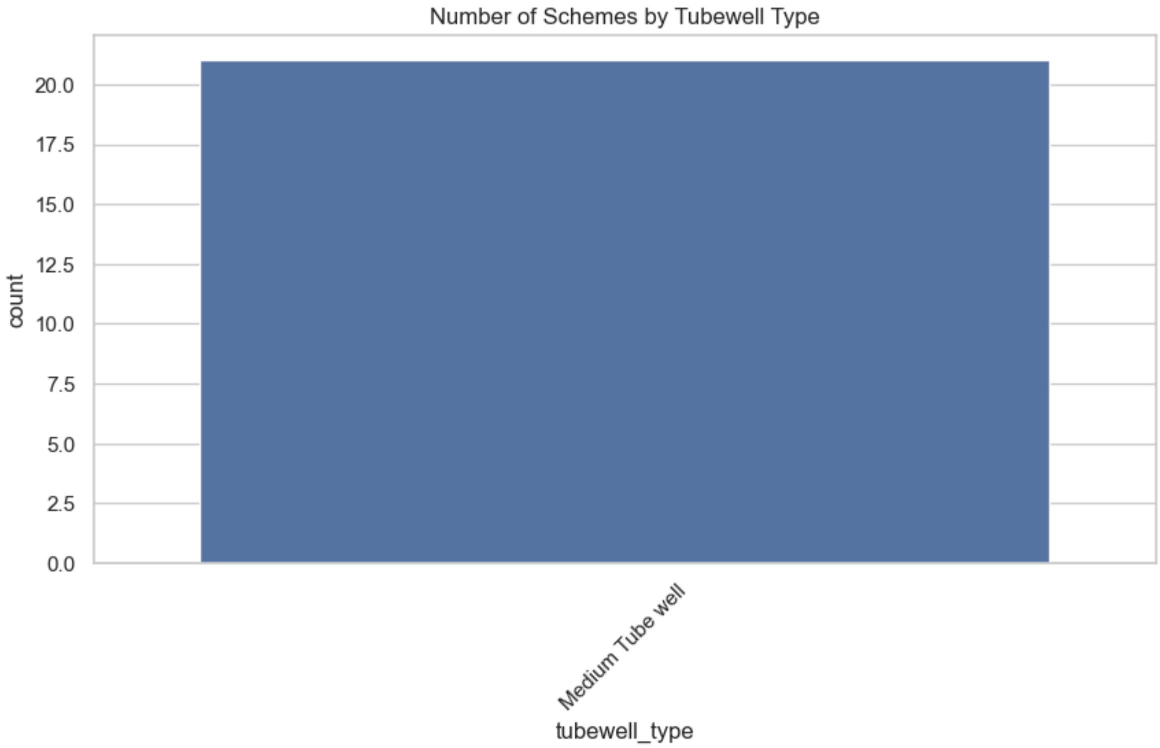
21 rows × 61 columns

In [7]:
```
sns.set(style="whitegrid")
```

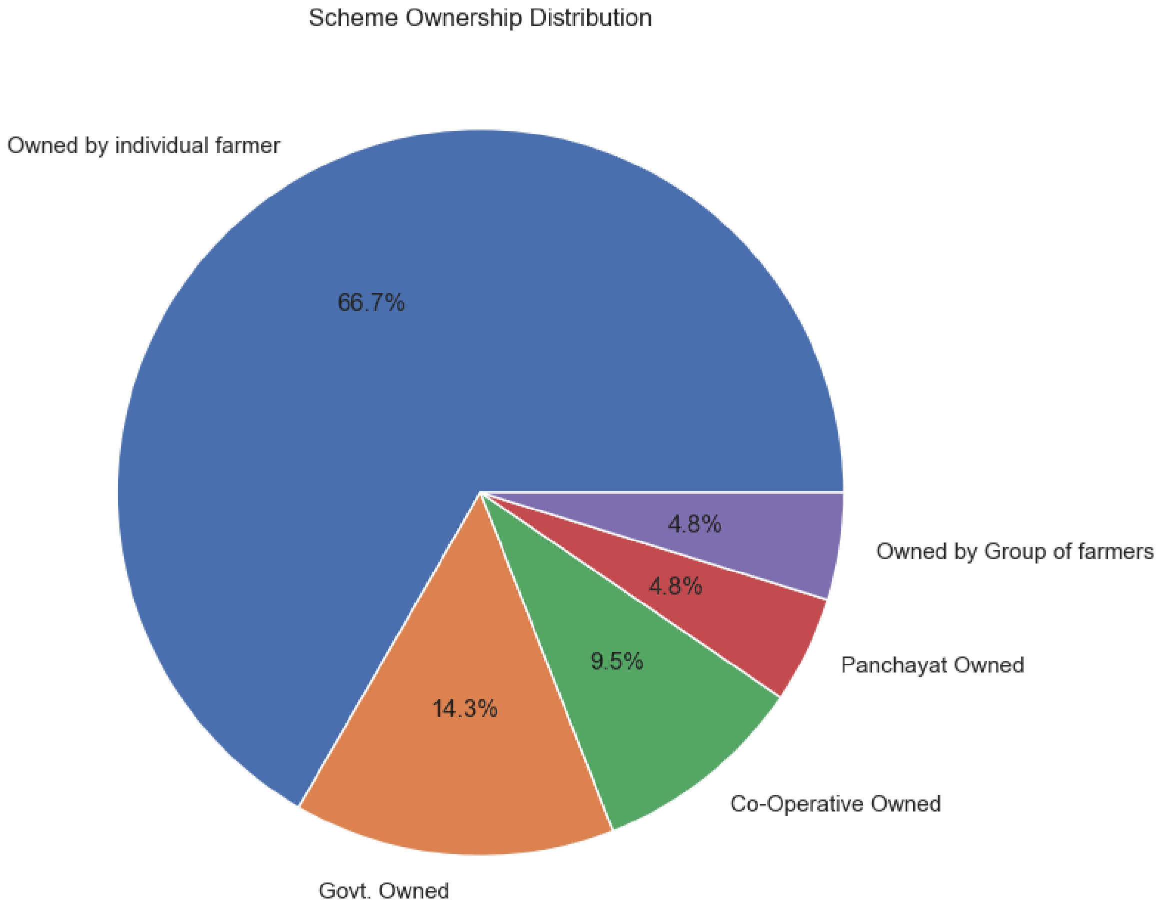
In [8]:
```
plt.figure(figsize=(10, 5))
sns.countplot(data=data, x='district_name')
plt.title('Number of Schemes by District')
plt.xticks(rotation=45)
plt.show()
```



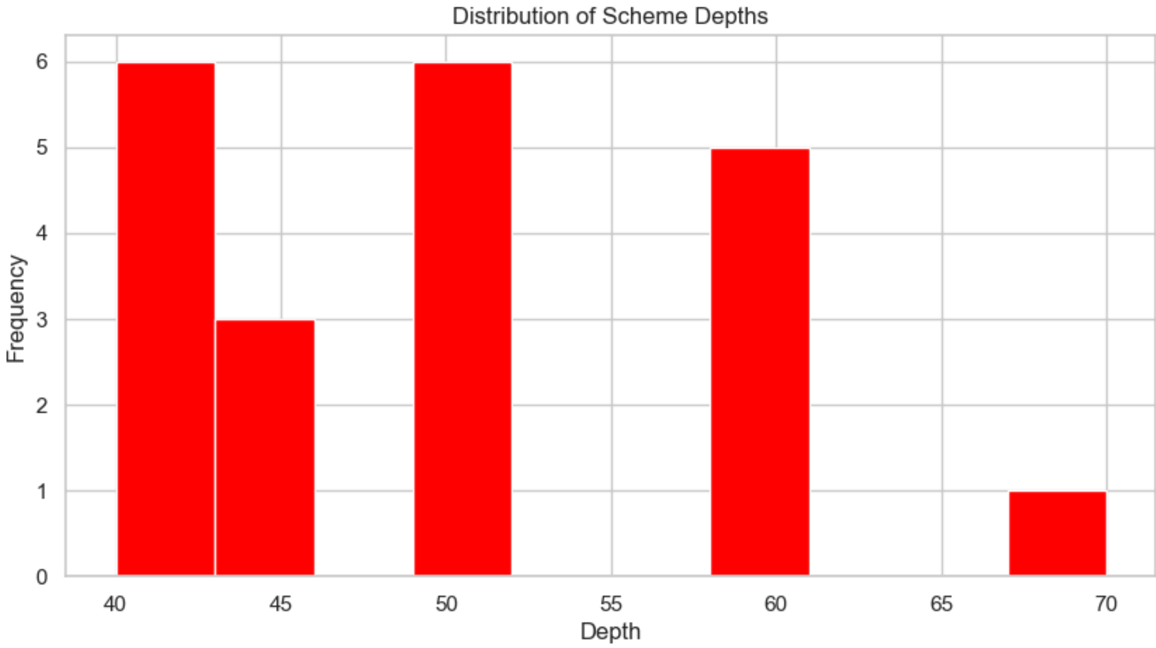
In [9]:
```
plt.figure(figsize=(10, 5))
sns.countplot(data=data, x='tubewell_type')
plt.title('Number of Schemes by Tubewell Type')
plt.xticks(rotation=45)
plt.show()
```



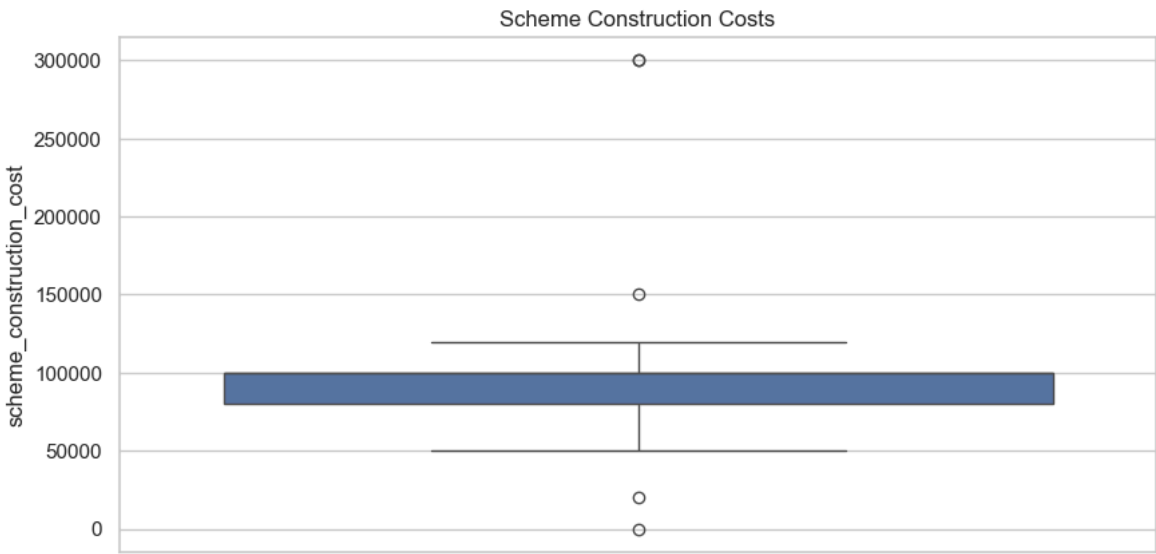
In [10]:
```
plt.figure(figsize=(8, 8))
data['scheme_owner_name'].value_counts().plot.pie(autopct='%1.1f%%')
plt.title('Scheme Ownership Distribution')
plt.ylabel('')
plt.show()
```



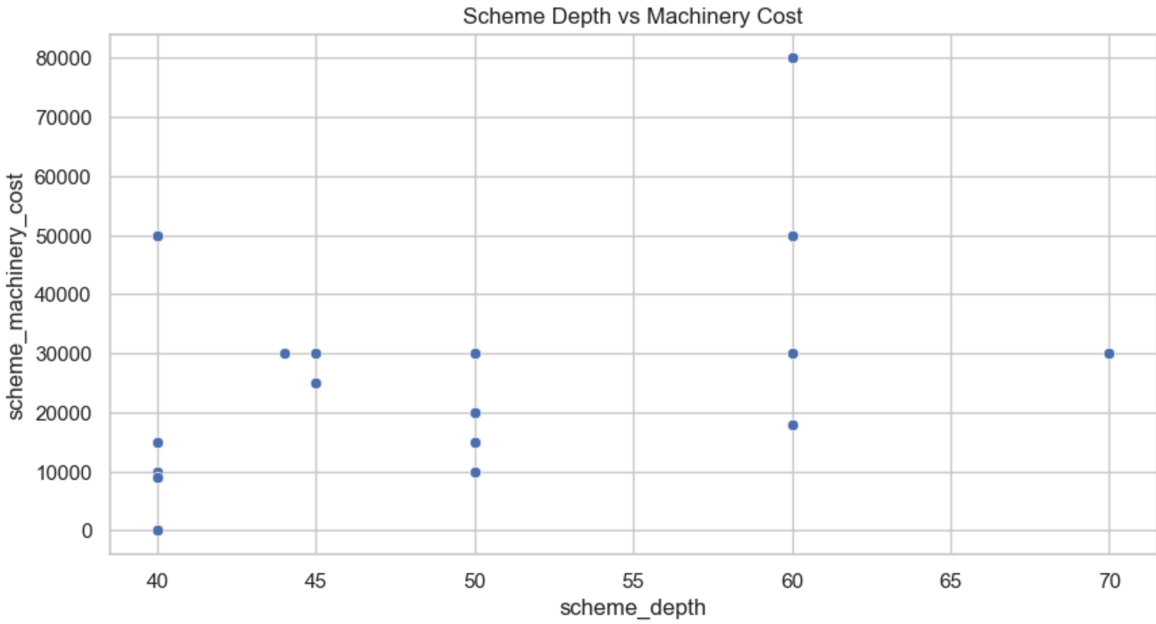
In [28]:
```
plt.figure(figsize=(10, 5))
plt.hist(data['scheme_depth'], bins=10, color='red')
plt.title('Distribution of Scheme Depths')
plt.xlabel('Depth')
plt.ylabel('Frequency')
plt.show()
```



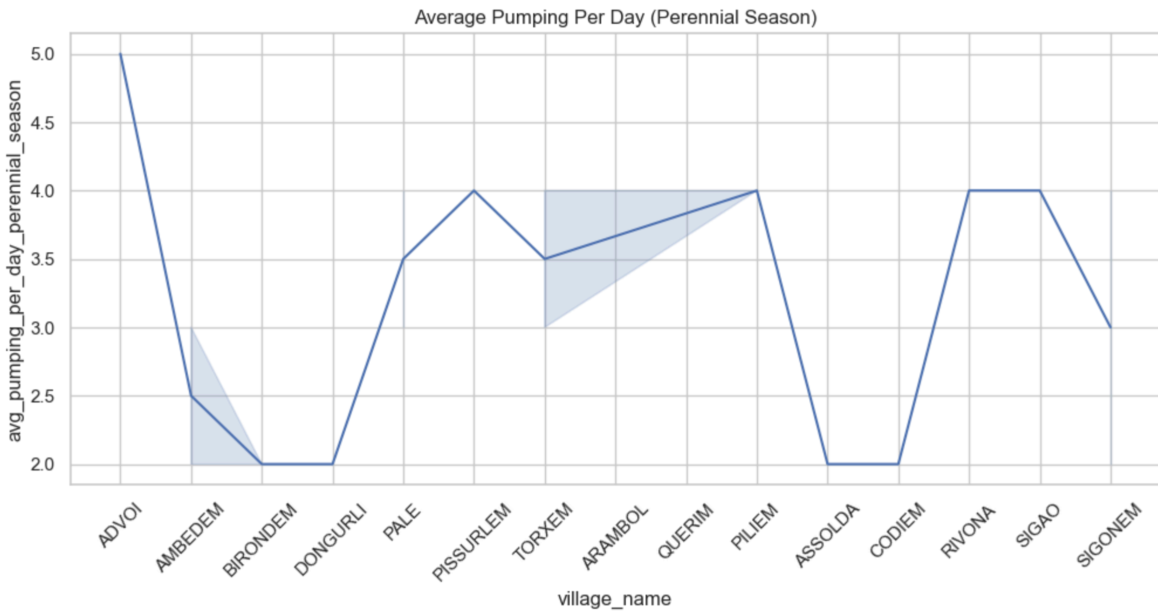
In [12]:
```
plt.figure(figsize=(10, 5))
sns.boxplot(data['scheme_construction_cost'])
plt.title('Scheme Construction Costs')
plt.show()
```



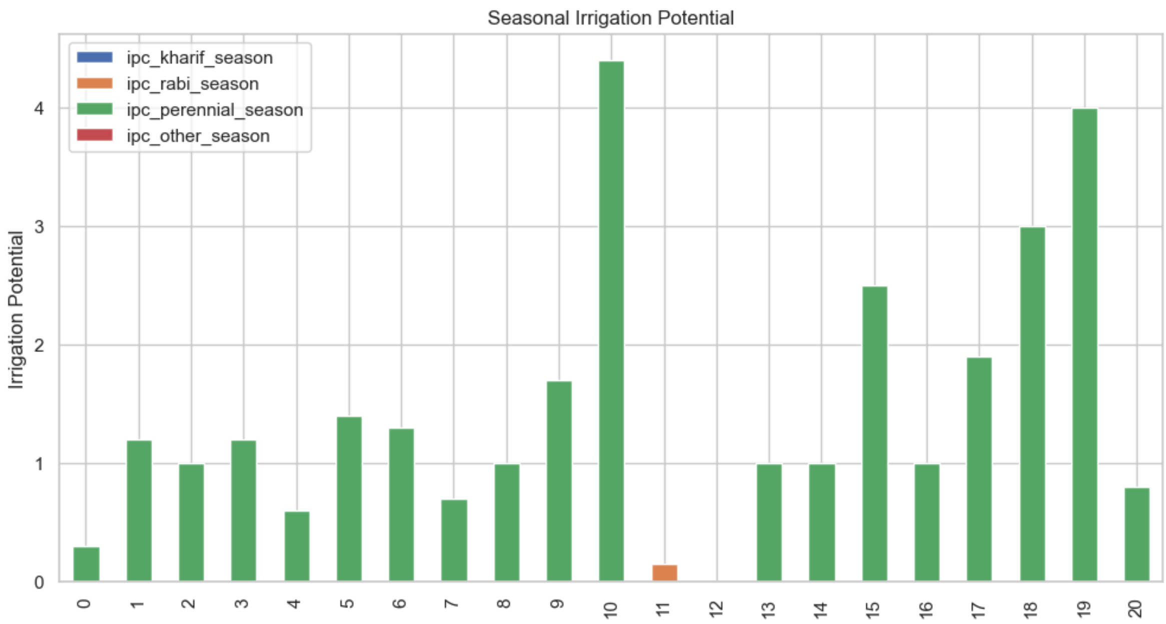
In [13]:
```
plt.figure(figsize=(10, 5))
sns.scatterplot(x='scheme_depth', y='scheme_machinery_cost', data=data)
plt.title('Scheme Depth vs Machinery Cost')
plt.show()
```



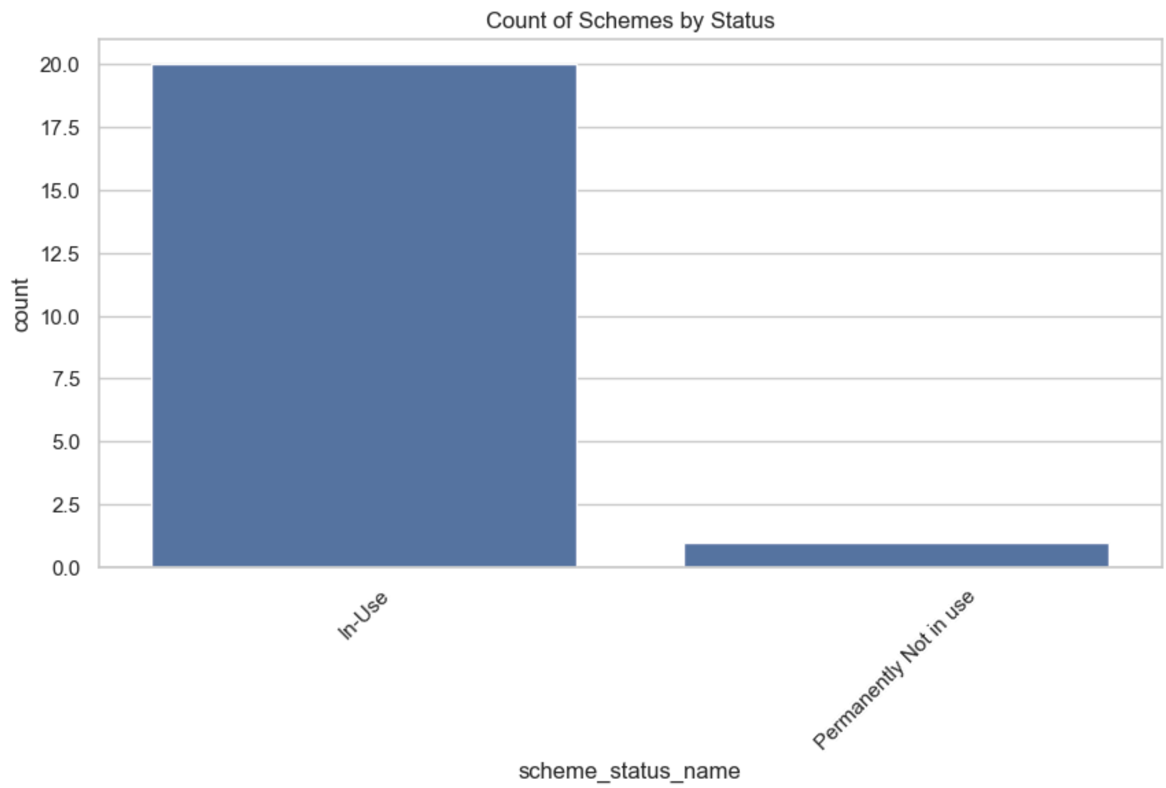
In [14]:
```
plt.figure(figsize=(12, 5))
sns.lineplot(x='village_name', y='avg_pumping_per_day_perennial_season', data=da
plt.xticks(rotation=45)
plt.title('Average Pumping Per Day (Perennial Season)')
plt.show()
```



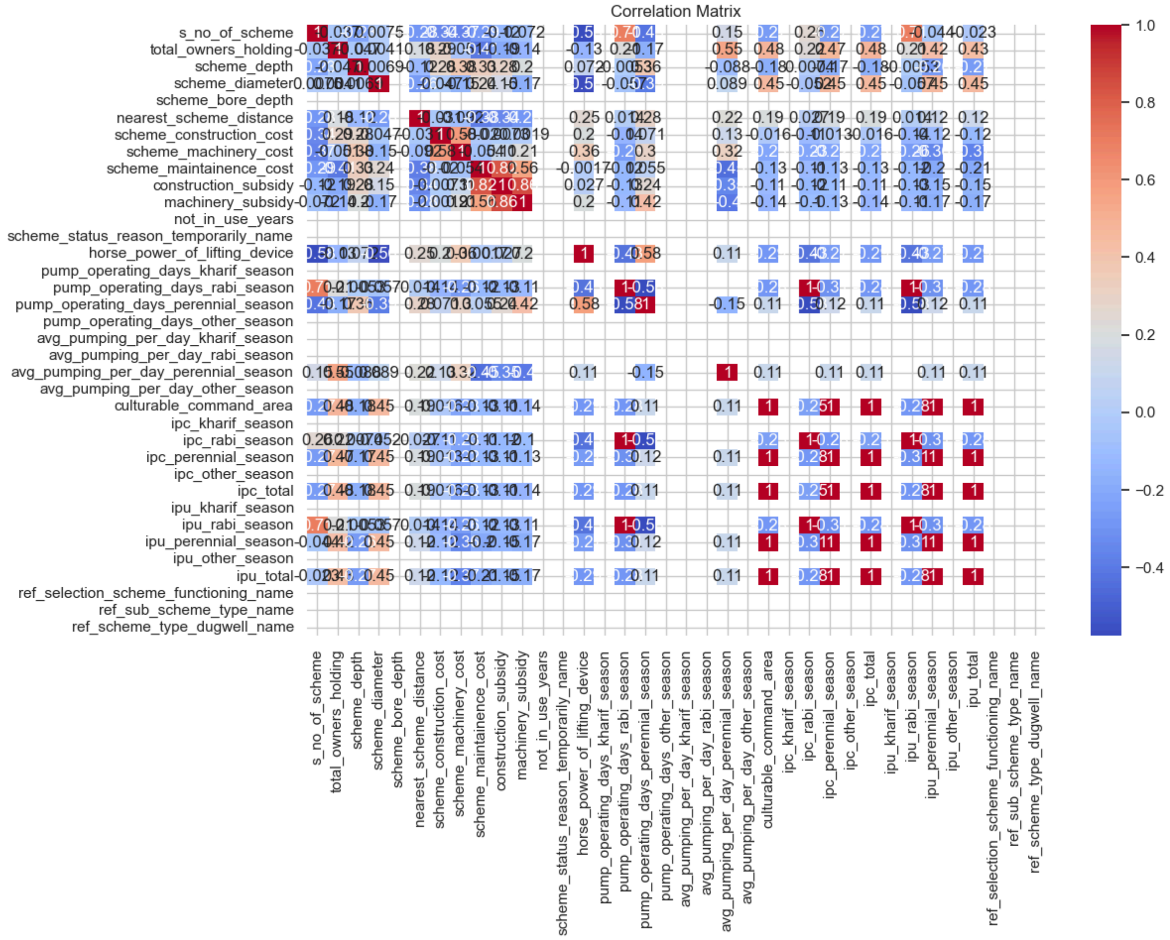
In [15]:
```
seasonal_cols = ['ipc_kharif_season', 'ipc_rabi_season', 'ipc_perennial_season',
data[seasonal_cols].plot(kind='bar', stacked=True, figsize=(12, 6))
plt.title('Seasonal Irrigation Potential')
plt.ylabel('Irrigation Potential')
plt.show()
```



In [16]:
```
plt.figure(figsize=(10, 5))
sns.countplot(data=data, x='scheme_status_name')
plt.title('Count of Schemes by Status')
plt.xticks(rotation=45)
plt.show()
```

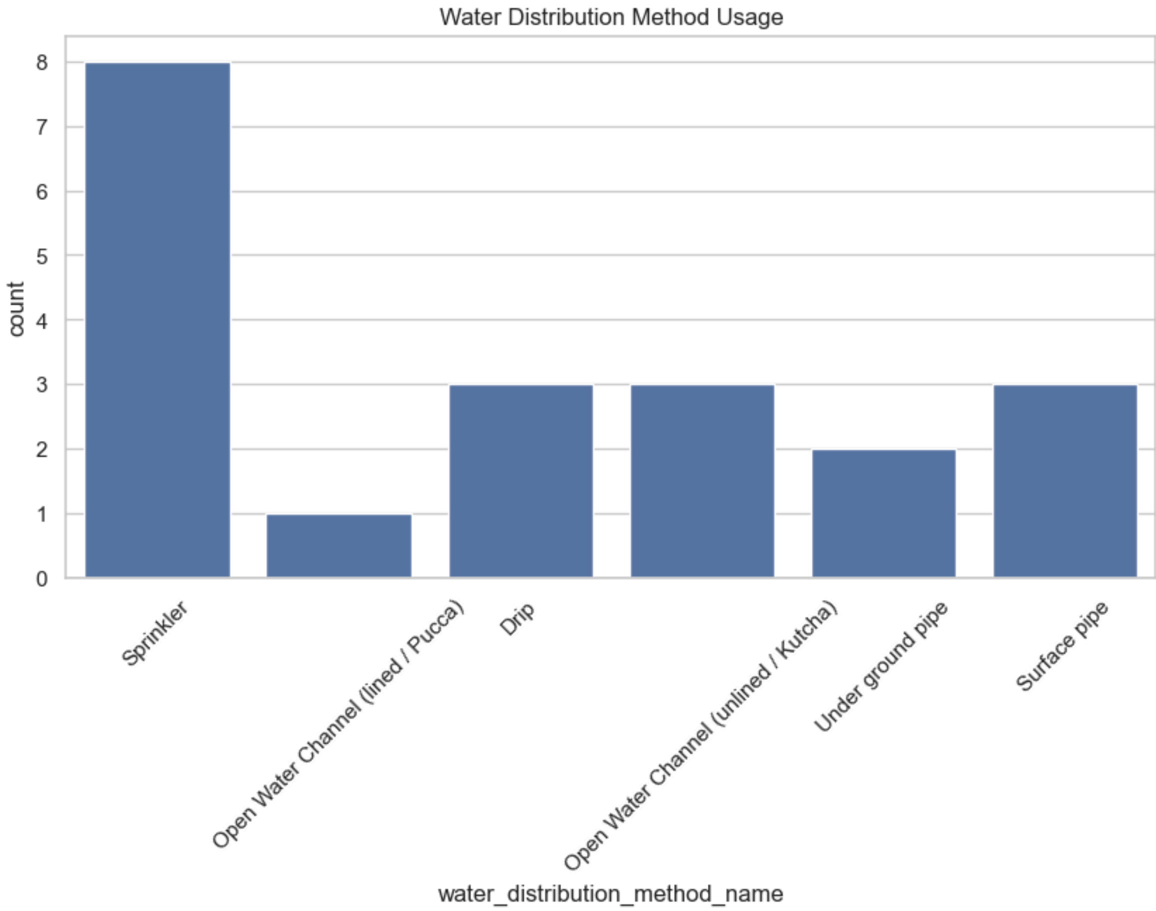


In [17]:
```
plt.figure(figsize=(12, 8))
sns.heatmap(data.corr(numeric_only=True), annot=True, cmap='coolwarm')
plt.title('Correlation Matrix')
plt.show()
```

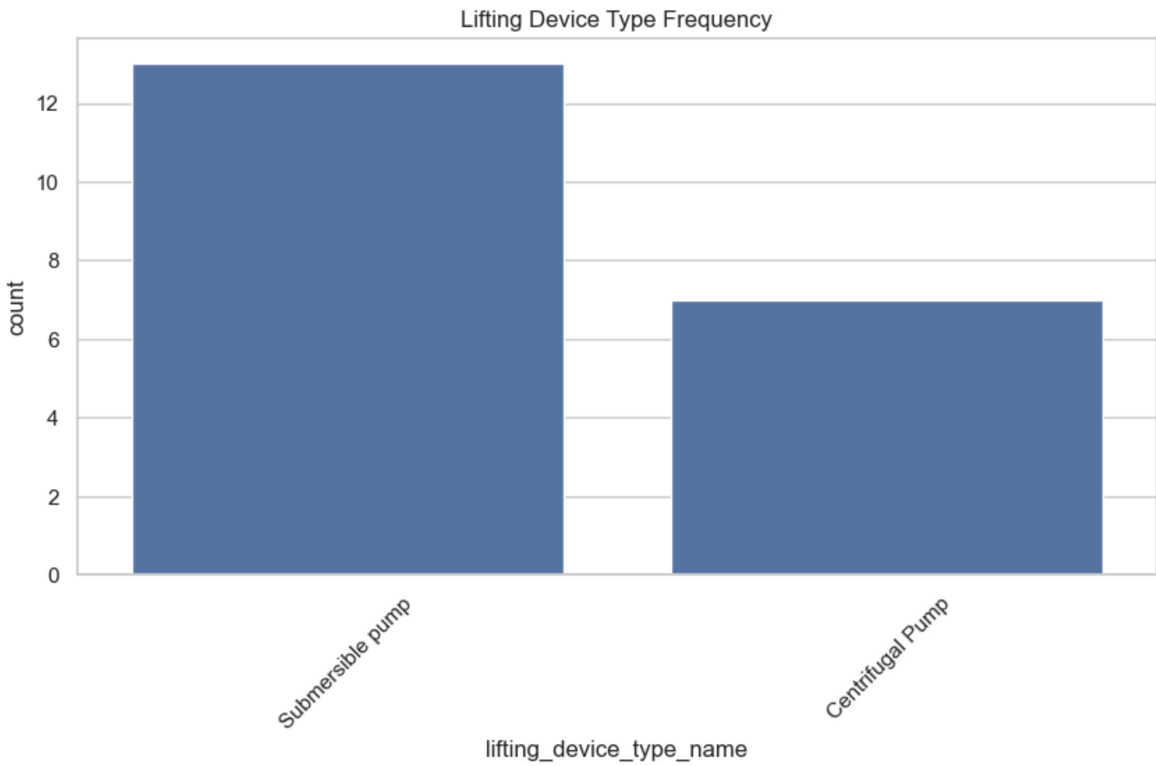


In [18]:

```
plt.figure(figsize=(10, 5))
sns.countplot(data=data, x='water_distribution_method_name')
plt.title('Water Distribution Method Usage')
plt.xticks(rotation=45)
plt.show()
```

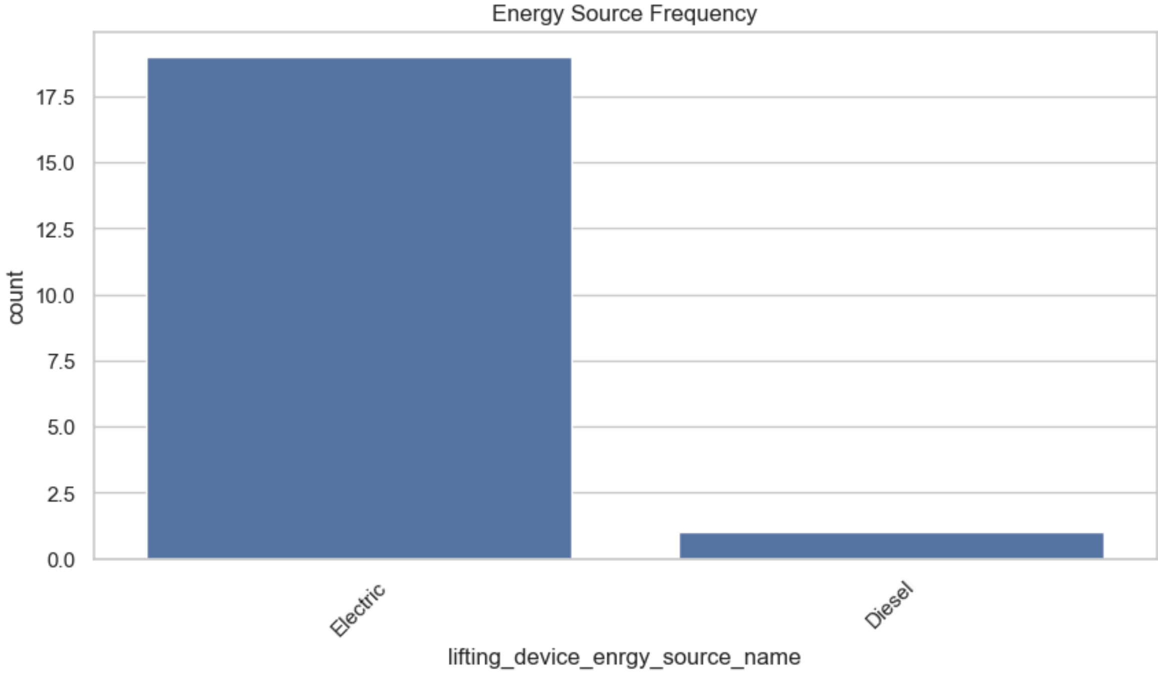


In [19]:
```
plt.figure(figsize=(10, 5))
sns.countplot(data=data, x='lifting_device_type_name')
plt.title('Lifting Device Type Frequency')
plt.xticks(rotation=45)
plt.show()
```

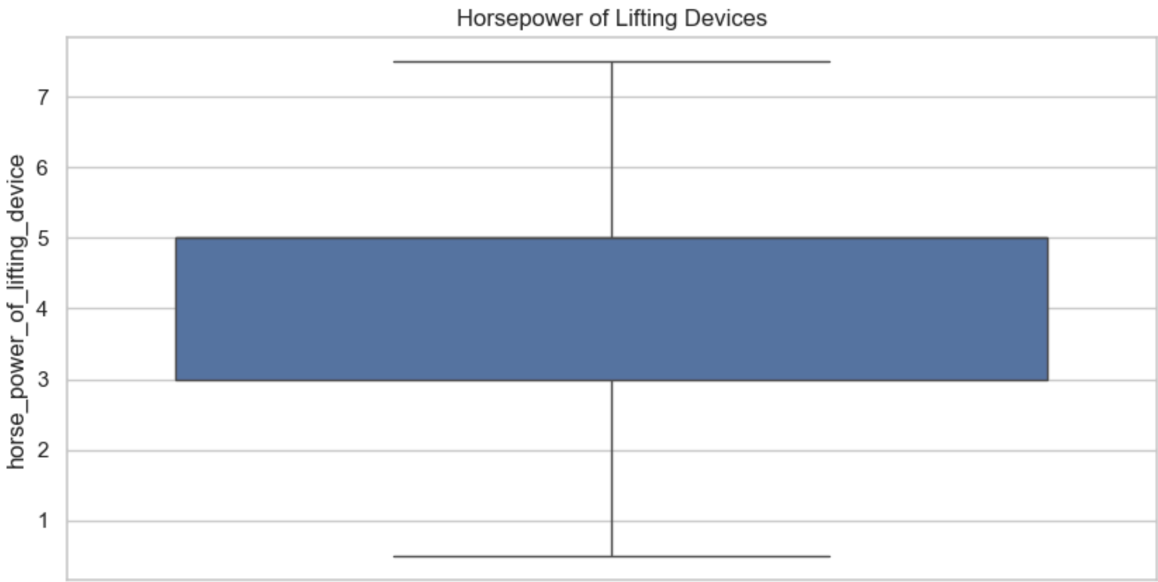


In [20]:
```
plt.figure(figsize=(10, 5))
sns.countplot(data=data, x='lifting_device_enrgy_source_name')
```

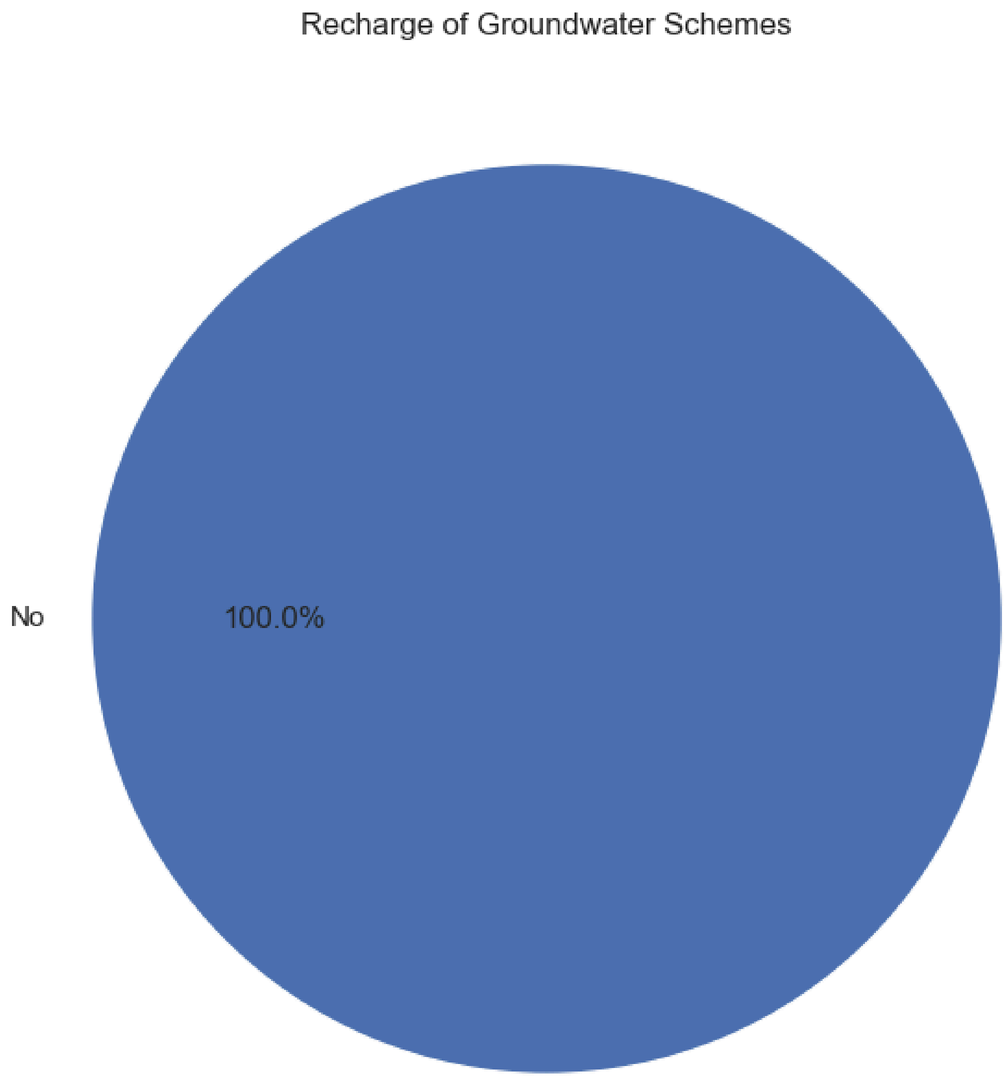
```
plt.title('Energy Source Frequency')
plt.xticks(rotation=45)
plt.show()
```



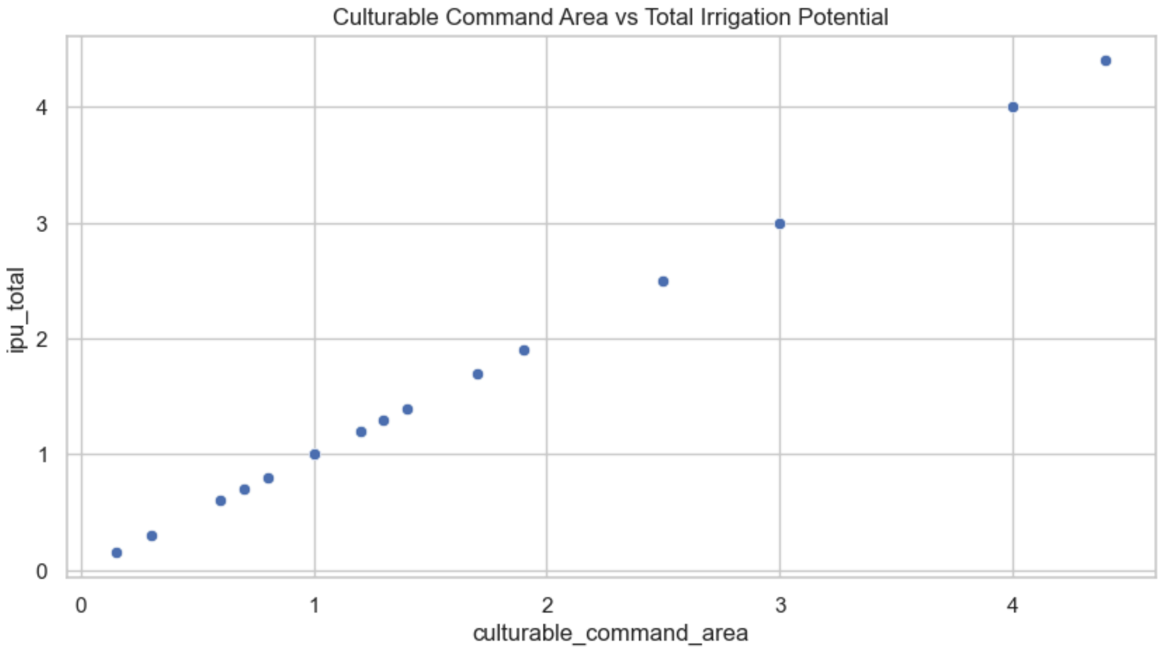
In [21]:
```
plt.figure(figsize=(10, 5))
sns.boxplot(data['horse_power_of_lifting_device'])
plt.title('Horsepower of Lifting Devices')
plt.show()
```



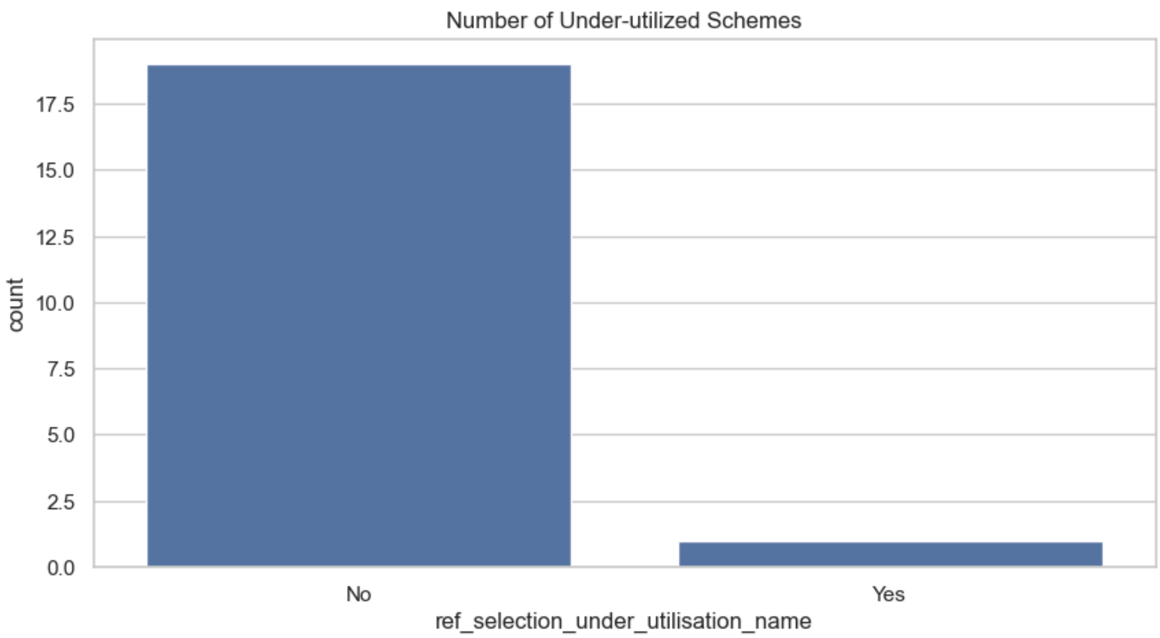
In [22]:
```
plt.figure(figsize=(8, 8))
data['ref_scheme_recharge_of_ground_water_name'].value_counts().plot.pie(autopct
plt.title('Recharge of Groundwater Schemes')
plt.ylabel('')
plt.show()
```



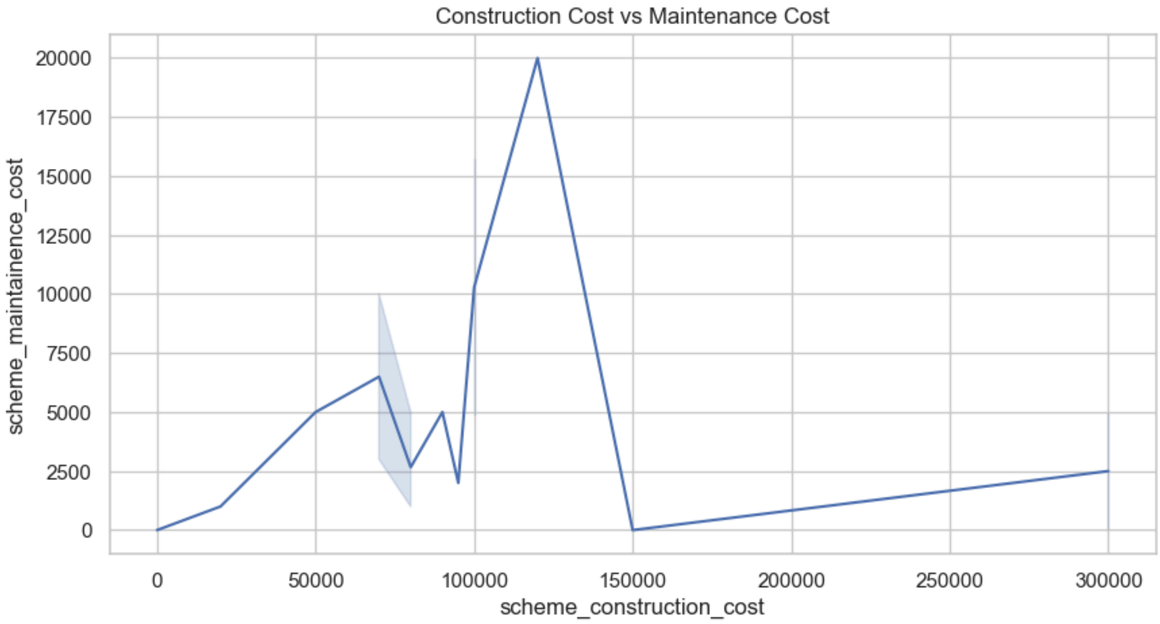
In [23]:
```
plt.figure(figsize=(10, 5))
sns.scatterplot(x='culturable_command_area', y='ipu_total', data=data)
plt.title('Culturable Command Area vs Total Irrigation Potential')
plt.show()
```



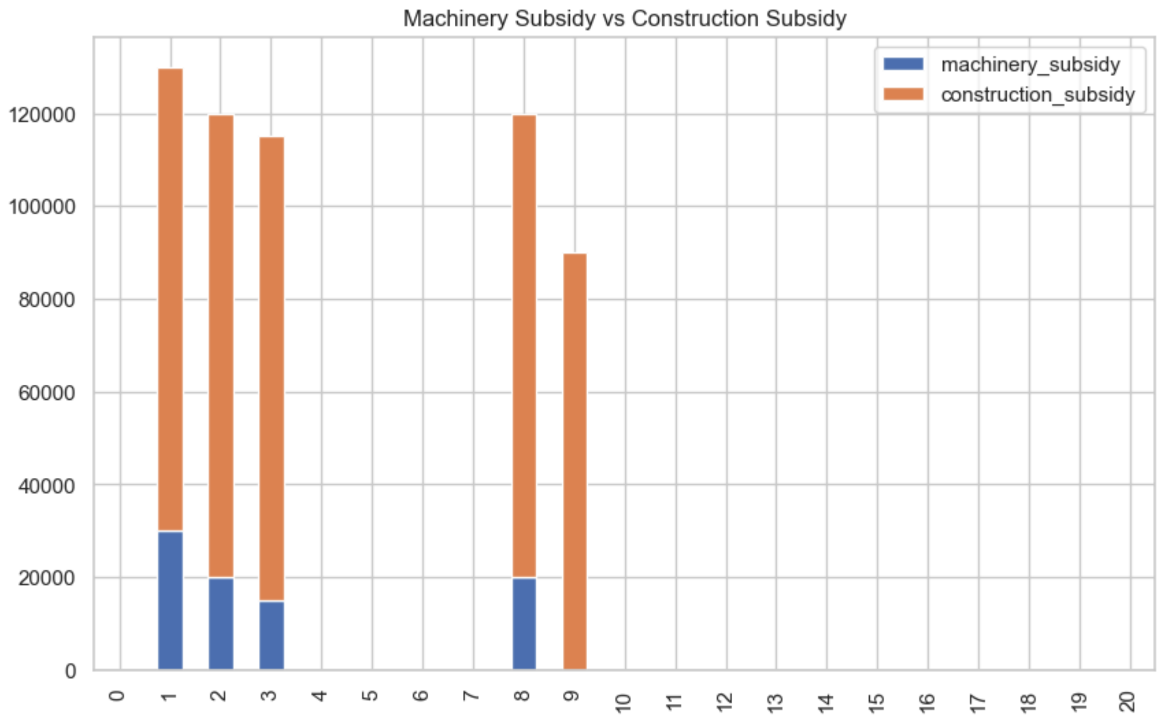
In [24]:
```
plt.figure(figsize=(10, 5))
sns.countplot(data=data, x='ref_selection_under_utilisation_name')
plt.title('Number of Under-utilized Schemes')
plt.show()
```



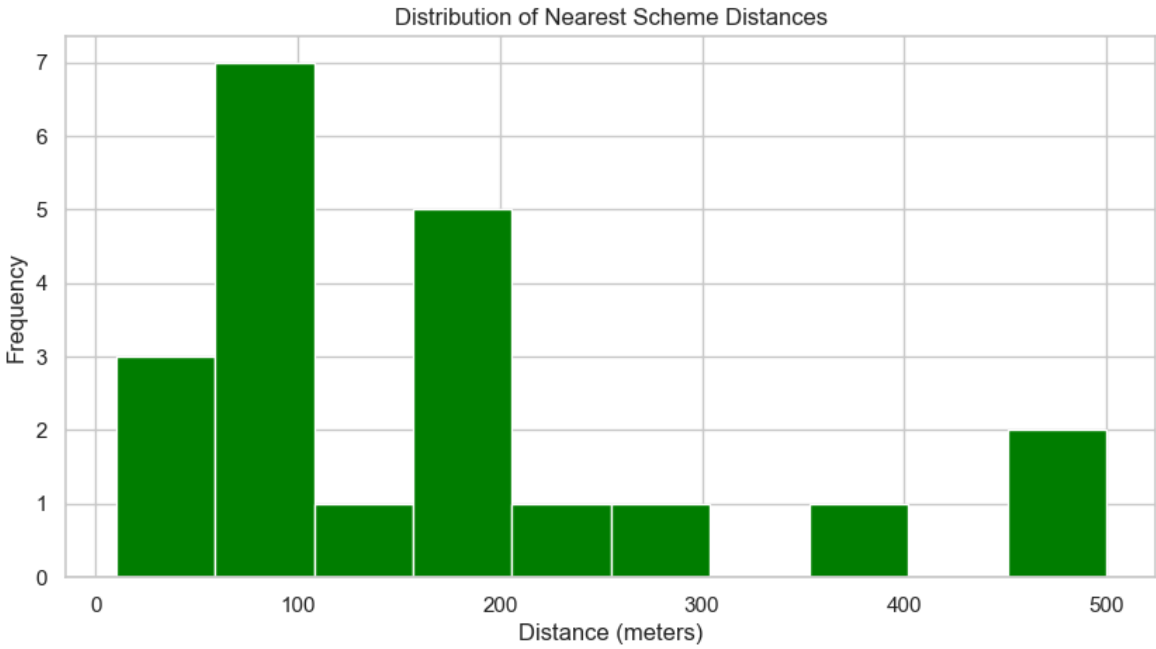
In [25]:
```
plt.figure(figsize=(10, 5))
sns.lineplot(x='scheme_construction_cost', y='scheme_maintainence_cost', data=da
plt.title('Construction Cost vs Maintenance Cost')
plt.show()
```



In [26]:
```
subsidy_cols = ['machinery_subsidy', 'construction_subsidy']
data[subsidy_cols].plot(kind='bar', stacked=True, figsize=(10, 6))
plt.title('Machinery Subsidy vs Construction Subsidy')
plt.show()
```



In [29]:
```
plt.figure(figsize=(10, 5))
plt.hist(data['nearest_scheme_distance'], bins=10, color='green')
plt.title('Distribution of Nearest Scheme Distances')
plt.xlabel('Distance (meters)')
plt.ylabel('Frequency')
plt.show()
```


Distribution of Nearest Scheme Distances
Frequency
Distance (meters)
0
1
2
3
4
5
6
7
0
100
200
300
400
500

```
In [ ]:
```