```
In [1]:  # Dependencies and Setup
         %matplotlib inline
         import matplotlib.pyplot as plt
         import pandas as pd
         import numpy as np

         # Hide warning messages in notebook
         import warnings
         warnings.filterwarnings('ignore')

         # File to Load (Remember to Change These)
         mouse_drug_data_to_load = "data/mouse_drug_data.csv"
         clinical_trial_data_to_load = "data/clinicaltrial_data.csv"

         # Read the Mouse and Drug Data and the Clinical Trial Data
         mouse_drug_data_df = pd.read_csv(mouse_drug_data_to_load)
         clinical_trial_data_df = pd.read_csv(clinical_trial_data_to_load)

         # Combine the data into a single dataset
         combined_mouse_clinical_df = pd.merge(clinical_trial_data_df,mouse_drug_data_d
         f, how='left', on='Mouse ID')

         # Display the data table for preview
         combined_mouse_clinical_df.head()
```

Out[1]:

	Mouse ID	Timepoint	Tumor Volume (mm3)	Metastatic Sites	Drug
0	b128	0	45.0	0	Capomulin
1	f932	0	45.0	0	Ketapril
2	g107	0	45.0	0	Ketapril
3	a457	0	45.0	0	Ketapril
4	c819	0	45.0	0	Ketapril

Tumor Response to Treatment

In [2]:
```python
# Store the Mean Tumor Volume Data Grouped by Drug and Timepoint
mean_tumor_volume = combined_mouse_clinical_df.groupby(['Drug','Timepoint'])[
'Tumor Volume (mm3)'].mean()

# Convert to DataFrame
mean_tumor_volume_df  = pd.DataFrame(mean_tumor_volume)
mean_tumor_volume_df

# Preview DataFrame
mean_tumor_volume_df=mean_tumor_volume_df.reset_index()
mean_tumor_volume_df.head()
```

Out[2]:

	Drug	Timepoint	Tumor Volume (mm3)
0	Capomulin	0	45.000000
1	Capomulin	5	44.266086
2	Capomulin	10	43.084291
3	Capomulin	15	42.064317
4	Capomulin	20	40.716325

In [3]:
```python
# Store the Standard Error of Tumor Volumes Grouped by Drug and Timepoint
mean_tumor_volume_ste = combined_mouse_clinical_df.groupby(['Drug','Timepoint'
])['Tumor Volume (mm3)'].sem()

# Convert to DataFrame

mean_tumor_volume_std_error = pd.DataFrame(mean_tumor_volume_ste).reset_index
()

# Preview DataFrame

mean_tumor_volume_std_error.head()
```

Out[3]:

	Drug	Timepoint	Tumor Volume (mm3)
0	Capomulin	0	0.000000
1	Capomulin	5	0.448593
2	Capomulin	10	0.702684
3	Capomulin	15	0.838617
4	Capomulin	20	0.909731

In [4]:
```python
# Minor Data Munging to Re-Format the Data Frames

mean_tumor_volume_pivot_df = mean_tumor_volume_df.pivot_table('Tumor Volume (mm3)', ['Timepoint'], 'Drug')[["Capomulin", "Infubinol", "Ketapril","Placebo"]]

# Preview that Reformatting worked

mean_tumor_volume_pivot_df
```

Out[4]:

Drug Timepoint	Capomulin	Infubinol	Ketapril	Placebo
0	45.000000	45.000000	45.000000	45.000000
5	44.266086	47.062001	47.389175	47.125589
10	43.084291	49.403909	49.582269	49.423329
15	42.064317	51.296397	52.399974	51.359742
20	40.716325	53.197691	54.920935	54.364417
25	39.939528	55.715252	57.678982	57.482574
30	38.769339	58.299397	60.994507	59.809063
35	37.816839	60.742461	63.371686	62.420615
40	36.958001	63.162824	66.068580	65.052675
45	36.236114	65.755562	70.662958	68.084082

In [5]:
```python
# Minor Data Munging to Re-Format the Data Frames

mean_tumor_volume_pivot_error_df=mean_tumor_volume_std_error.pivot_table('Tumor Volume (mm3)', ['Timepoint'], 'Drug')[["Capomulin", "Infubinol", "Ketapril", "Placebo"]]


# Preview that Reformatting worked

mean_tumor_volume_pivot_error_df.head()
```

Out[5]:

Drug Timepoint	Capomulin	Infubinol	Ketapril	Placebo
0	0.000000	0.000000	0.000000	0.000000
5	0.448593	0.235102	0.264819	0.218091
10	0.702684	0.282346	0.357421	0.402064
15	0.838617	0.357705	0.580268	0.614461
20	0.909731	0.476210	0.726484	0.839609

In [6]:

```python
# Generate the Plot (with Error Bars)

plt.errorbar(mean_tumor_volume_pivot_df.index, mean_tumor_volume_pivot_df['Cap
omulin'],
             yerr=mean_tumor_volume_pivot_error_df['Capomulin'],
             color='red', marker='+', markersize=5, linestyle='-', linewidth=1
,label='Capomulin')
plt.errorbar(mean_tumor_volume_pivot_df.index, mean_tumor_volume_pivot_df['Inf
ubinol'],
             yerr=mean_tumor_volume_pivot_error_df['Infubinol'],
             color='blue', marker='o', markersize=5, linestyle='-', linewidth=
1,label='Infubinol')
plt.errorbar(mean_tumor_volume_pivot_df.index, mean_tumor_volume_pivot_df['Ket
april'],
             yerr=mean_tumor_volume_pivot_error_df['Ketapril'],
             color='green', marker='x', markersize=5, linestyle='-', linewidth
=1,label='Ketapril')
plt.errorbar(mean_tumor_volume_pivot_df.index, mean_tumor_volume_pivot_df['Pla
cebo'],
             yerr=mean_tumor_volume_pivot_error_df['Placebo'],
             color='black', marker='^', markersize=5, linestyle='-', linewidth
=1,label='Placebo')

# Create Labels for the X and Y axis

plt.xlabel("Time(Days)")
plt.ylabel("Tumor Volume (mm3)")
plt.title("Tumor Response to Treatment")
plt.xlim(0,50)
plt.ylim(35,75)
plt.grid()

# Set our legend to where the chart thinks is best
plt.legend(loc='best')


# Save the Figure
plt.savefig("../Pymaceuticals/Images/tumor_response.png")

# Show the Figure
plt.show()
```
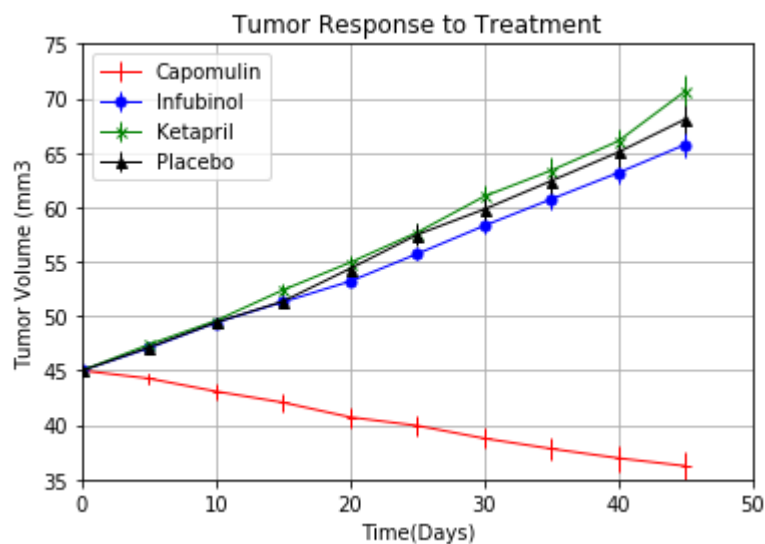
Tumor Response to Treatment

Tumor Response to Treatment

Metastatic Response to Treatment

```
In [7]:  # Store the Mean Met. Site Data Grouped by Drug and Timepoint

         mean_met_site = combined_mouse_clinical_df.groupby(['Drug','Timepoint'])['Meta
         static Sites'].mean()
         mean_met_site

         # Convert to DataFrame

         mean_met_site_df = pd.DataFrame(mean_met_site)

         # Preview DataFrame

         mean_met_site_df.head()
```

Out[7]:

		Metastatic Sites
Drug	**Timepoint**	
	0	0.000000
	5	0.160000
Capomulin	10	0.320000
	15	0.375000
	20	0.652174

In [8]:
```python
# Store the Standard Error associated with Met. Sites Grouped by Drug and Time
point

mean_met_site_err = combined_mouse_clinical_df.groupby(['Drug','Timepoint'])[
'Metastatic Sites'].sem()
mean_met_site_err


# Convert to DataFrame

mean_met_site_err_df = pd.DataFrame(mean_met_site_err)

# Preview DataFrame

mean_met_site_err_df
```

Out[8]:

		Metastatic Sites
Drug	**Timepoint**	
	0	0.000000
	5	0.074833
Capomulin	10	0.125433
	15	0.132048
	20	0.161621
...	...	...
	25	0.236621
	30	0.248168
Zoniferol	35	0.285714
	40	0.299791
	45	0.286400

100 rows × 1 columns

In [9]:
```python
# Minor Data Munging to Re-Format the Data Frames

rows_to_column_met_pivot = mean_met_site_df.pivot_table('Metastatic Sites', [
'Timepoint'], 'Drug')[["Capomulin", "Infubinol", "Ketapril","Placebo"]]
mean_met_site_pivot_df = pd.DataFrame(rows_to_column_met_pivot)


# Preview that Reformatting worked
mean_met_site_pivot_df.head()
```

Out[9]:

Drug	Capomulin	Infubinol	Ketapril	Placebo
Timepoint				
0	0.000000	0.000000	0.000000	0.000000
5	0.160000	0.280000	0.304348	0.375000
10	0.320000	0.666667	0.590909	0.833333
15	0.375000	0.904762	0.842105	1.250000
20	0.652174	1.050000	1.210526	1.526316

In [10]:
```python
# Minor Data Munging to Re-Format the Data Frames

rows_to_column_met_err_pivot = mean_met_site_err_df.pivot_table('Metastatic Si
tes', ['Timepoint'], 'Drug')[["Capomulin", "Infubinol", "Ketapril","Placebo"]]
mean_met_site_pivot_err_df = pd.DataFrame(rows_to_column_met_err_pivot)


# Preview that Reformatting worked
mean_met_site_pivot_err_df.head()
```

Out[10]:

Drug	Capomulin	Infubinol	Ketapril	Placebo
Timepoint				
0	0.000000	0.000000	0.000000	0.000000
5	0.074833	0.091652	0.098100	0.100947
10	0.125433	0.159364	0.142018	0.115261
15	0.132048	0.194015	0.191381	0.190221
20	0.161621	0.234801	0.236680	0.234064

In [11]:
```python
Generate the Plot (with Error Bars)

plt.errorbar(mean_met_site_pivot_df.index, mean_met_site_pivot_df['Capomulin'
], yerr=mean_met_site_pivot_err_df['Capomulin'],
            color='red', marker='+', markersize=5, linestyle='-', linewidth=1
,label='Capomulin')
plt.errorbar(mean_met_site_pivot_df.index, mean_met_site_pivot_df['Infubinol'
], yerr=mean_met_site_pivot_err_df['Infubinol'],
            color='blue', marker='o', markersize=5, linestyle='-', linewidth=
1,label='Infubinol')
plt.errorbar(mean_met_site_pivot_df.index, mean_met_site_pivot_df['Ketapril'],
yerr=mean_met_site_pivot_err_df['Ketapril'],
            color='green', marker='x', markersize=5, linestyle='-', linewidth
=1,label='Ketapril')
plt.errorbar(mean_met_site_pivot_df.index, mean_met_site_pivot_df['Placebo'],
yerr=mean_met_site_pivot_err_df['Placebo'],
            color='black', marker='^', markersize=5, linestyle='-', linewidth
=1,label='Placebo')

Create Labels,Limits for the X and Y axis

plt.xlabel("Treatment Duration (Days)")
plt.ylabel("Met.Sites")
plt.title("Matastatic Spread During Treatment")
plt.xlim(0,50)
plt.ylim(0,4)
plt.grid()
plt.legend(loc="best")

Save the Figure

plt.savefig("../Pymaceuticals/Images/Matastatic_Spread.png")
# Show the Figure
plt.show()
```
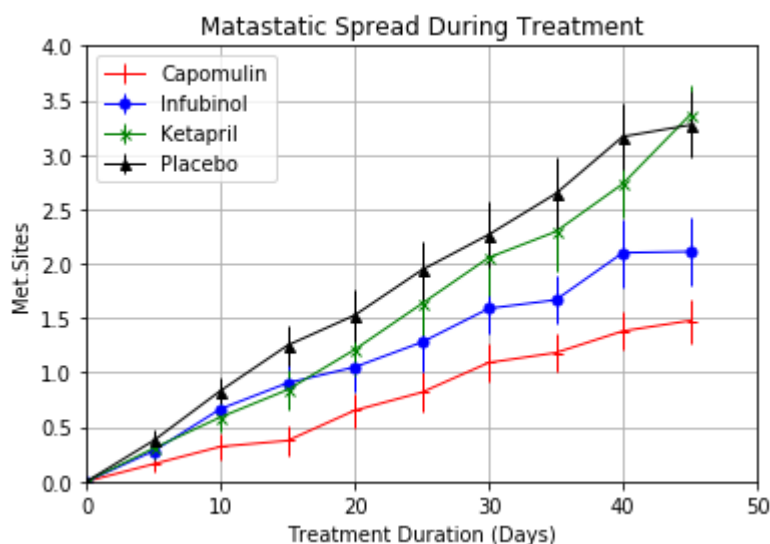


Metastatic Spread During Treatment

# Survival Rates

In [12]:
```python
# Store the Count of Mice Grouped by Drug and Timepoint (W can pass any metric)

mice_count = combined_mouse_clinical_df.groupby(['Drug','Timepoint'])['Mouse ID'].count()
mice_count

# Convert to DataFrame

mice_count_df = pd.DataFrame(mice_count).reset_index()

# Preview DataFrame

mice_count_df.head()
```

Out[12]:

	Drug	Timepoint	Mouse ID
0	Capomulin	0	25
1	Capomulin	5	25
2	Capomulin	10	25
3	Capomulin	15	24
4	Capomulin	20	23

In [13]:
```python
# Minor Data Munging to Re-Format the Data Frames

rows_to_column_mice_pivot = mice_count_df.pivot_table('Mouse ID', ['Timepoint'
], 'Drug')\
[["Capomulin", "Infubinol", "Ketapril","Placebo"]]

mice_count_format_df = pd.DataFrame(rows_to_column_mice_pivot)


# Preview the Data Frame
mice_count_format_df
```

Out[13]:

Drug	Capomulin	Infubinol	Ketapril	Placebo
Timepoint				
0	25	25	25	25
5	25	25	23	24
10	25	21	22	24
15	24	21	19	20
20	23	20	19	19
25	22	18	19	17
30	22	17	18	15
35	22	12	17	14
40	21	10	15	12
45	21	9	11	11

In [14]:
```
mice_count_format_first= mice_count_format_df.iloc[0,1]

# Convert mice_count_format_df to percentage
percentage_mice_count= (mice_count_format_df/mice_count_format_first)*100
percentage_mice_count
```

Out[14]:

Drug Timepoint	Capomulin	Infubinol	Ketapril	Placebo
0	100.0	100.0	100.0	100.0
5	100.0	100.0	92.0	96.0
10	100.0	84.0	88.0	96.0
15	96.0	84.0	76.0	80.0
20	92.0	80.0	76.0	76.0
25	88.0	72.0	76.0	68.0
30	88.0	68.0	72.0	60.0
35	88.0	48.0	68.0	56.0
40	84.0	40.0	60.0	48.0
45	84.0	36.0	44.0	44.0

In [15]:
```python
# Generate the Plot (Accounting for percentages)

plt.errorbar(percentage_mice_count.index, percentage_mice_count['Capomulin'],
             color='red', marker='+', markersize=5, linestyle='-', linewidth=1
,label='Capomulin')

plt.errorbar(percentage_mice_count.index, percentage_mice_count['Infubinol'],
             color='blue', marker='o', markersize=5, linestyle='-', linewidth=
1,label='Infubinol')

plt.errorbar(percentage_mice_count.index, percentage_mice_count['Ketapril'],
             color='green', marker='x', markersize=5, linestyle='-', linewidth
=1,label='Ketapril')

plt.errorbar(percentage_mice_count.index, percentage_mice_count['Placebo'],
             color='black', marker='^', markersize=5, linestyle='-', linewidth
=1,label='Placebo')

# Create Labels,Limits for the X and Y axis, Chart title

plt.title("Survival During Treatment")
plt.xlabel("Treatment Duration (Days)")
plt.ylabel("Survival Rate (%)")
plt.grid()
plt.legend(loc='best')

# Save the Figure
plt.savefig("../Pymaceuticals/Images/survival_treatment.png")

# Show the Figure
plt.show()
```
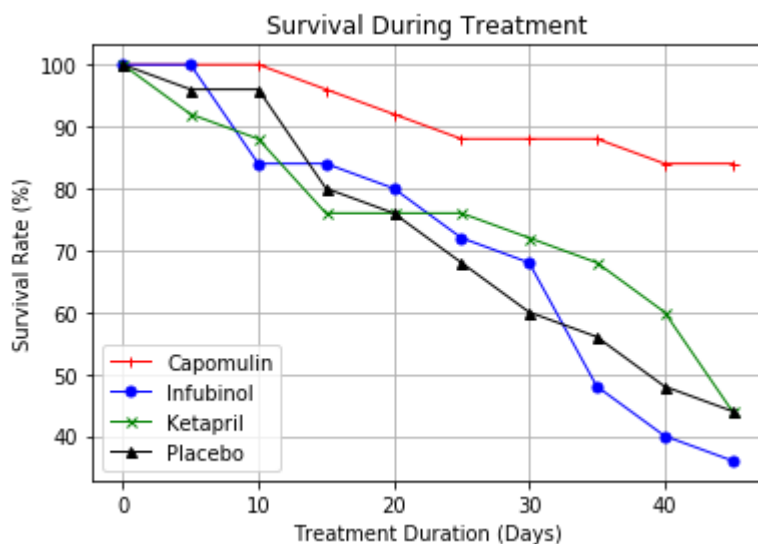

Metastatic Spread During Treatment

Summary Bar Graph

In [16]:
```python
# Calculate the percent changes for each drug
change_tumor_volume_first= mean_tumor_volume_pivot_df.iloc[0,:]
#print(change_tumor_volume_first)
change_tumor_volume_end= mean_tumor_volume_pivot_df.iloc[-1,:]
#print(change_tumor_volume_end)


percentage_change_tumor_volume= ((change_tumor_volume_end-change_tumor_volume_
first)/change_tumor_volume_first)*100


# Display the data to confirm
round(percentage_change_tumor_volume,5)
```

Out[16]:
```
Drug
Capomulin    -19.47530
Infubinol     46.12347
Ketapril      57.02879
Placebo       51.29796
dtype: float64
```

In [17]:
```python
percentage_change_tumor_volume.index
drug=tuple(percentage_change_tumor_volume.index)
```

In [18]:
```python
value = list(percentage_change_tumor_volume)
value
value2=[]
[ ]
for x in value:
    number=round(x,2)
    value2.append(number)
value2
value3 =tuple(value2)
value3
```

Out[18]: (-19.48, 46.12, 57.03, 51.3)

In [19]:
```python
# Store all Relevant Percent Changes into a Tuple
# Helped from TA in order to get the values in the graph

print(drug)
print(value3)


# Splice the data between passing and failing drugs

plt.bar(drug, value3, color=["green","red","red","red"], align="center")


# Orient widths. Add Labels, tick marks, etc.
tick_locations = [value for value in drug]
plt.xticks(tick_locations, drug)
plt.xlim(-0.50, len(drug)-0.50)
plt.ylim(-25, max(value)+5.0)

plt.title("Tumor Change Over 45 Day Treatment")
plt.ylabel("% Tumor Volume Change")
plt.grid()


# Use functions to label the percentages of changes
for a,b in zip(drug, value3):
    plt.text(a, b+1, str(b))

# Call functions to implement the function calls


# Save the Figure

plt.savefig("../Pymaceuticals/Images/tumor_change.png")
# Show the Figure
plt.show()
```
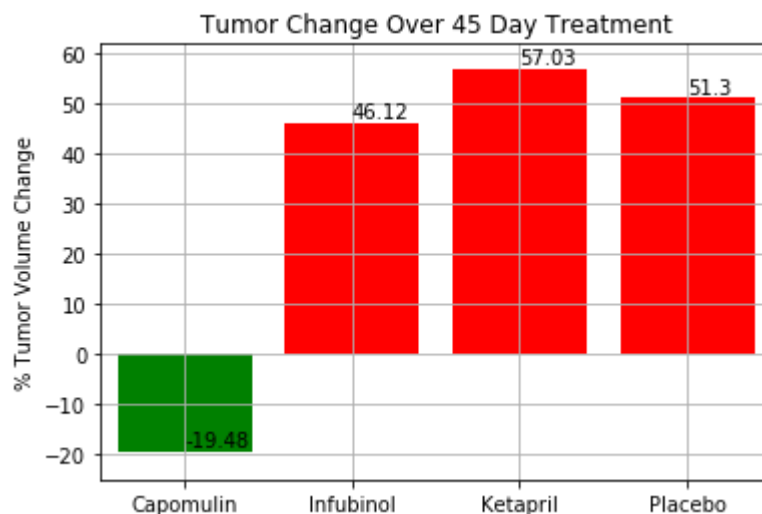
('Capomulin', 'Infubinol', 'Ketapril', 'Placebo')
(-19.48, 46.12, 57.03, 51.3)

Metastatic Spread During Treatment