

```
In [7]: # Dependencies and Setup
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np
import os
# Hide warning messages in notebook
import warnings
warnings.filterwarnings('ignore')

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

# File to Load (Remember to Change These)
mouse_data_path = os.path.join('data', 'mouse_drug_data.csv')
clinical_data_path = os.path.join('data', 'clinicaltrial_data.csv')

# Read the Mouse and Drug Data and the Clinical Trial Data
mouse_drug_df = pd.read_csv(mouse_data_path)
clinical_trial_df = pd.read_csv(clinical_data_path)

# Combine the data into a single dataset
complete_df = pd.merge(clinical_trial_df, mouse_drug_df, on='Mouse ID', how='left')

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

Out[7]:

	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 [3]: # Store the Mean Tumor Volume Data Grouped by Drug and Timepoint
grouped_df = complete_df.groupby(['Drug', 'Timepoint'])['Tumor Volume (mm3)']
avg_tumor_df = grouped_df.mean()

# Convert to DataFrame
avg_tumor_df = avg_tumor_df.reset_index()

# Preview DataFrame
avg_tumor_df.head()
```

Out[3]:

	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 [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
5	Capomulin	25	39.939528
6	Capomulin	30	38.769339
7	Capomulin	35	37.816839
8	Capomulin	40	36.958001
9	Capomulin	45	36.236114
10	Ceftamin	0	45.000000

```
In [4]: # Store the Standard Error of Tumor Volumes Grouped by Drug and Timepoint
tumor_standard_errors = grouped_df.sem()

# Convert to DataFrame
tumor_standard_errors_df = pd.DataFrame(tumor_standard_errors)

# Preview DataFrame
tumor_standard_errors_df.reset_index(inplace=True)
tumor_standard_errors_df.head()
```

Out[4]:

	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 [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 [5]: # Minor Data Munging to Re-Format the Data Frames
pivot_avg_tumor_df = avg_tumor_df.pivot(index='Timepoint',columns='Drug',values='Tumor Volume (mm3)')

# Preview that Reformating worked
pivot_avg_tumor_df.head(20)
```

```
Out[5]:
```

Timepoint	Drug	Capomulin	Ceftamin	Infubinol	Ketapril	Naftisol	Placebo	Propriiva	Ramicane	Stelasyn	Zoniferol
0	45.000000	45.000000	45.000000	45.000000	45.000000	45.000000	45.000000	45.000000	45.000000	45.000000	45.000000
5	44.266086	46.503051	47.062001	47.389175	46.796098	47.125589	47.248967	43.944859	47.527452	46.851818	
10	43.084291	48.285125	49.403909	49.582269	48.694210	49.423329	49.101541	42.531957	49.463844	48.689881	
15	42.064317	50.094055	51.296397	52.399974	50.933018	51.359742	51.067318	41.495061	51.529409	50.779059	
20	40.716325	52.157049	53.197691	54.920935	53.644087	54.364417	53.346737	40.238325	54.067395	53.170334	
25	39.939528	54.287674	55.715252	57.678982	56.731968	57.482574	55.504138	38.974300	56.166123	55.432935	
30	38.769334	56.769517	58.299397	60.994507	59.559509	59.809063	58.198374	38.703131	59.826738	57.713531	
35	37.816839	58.827548	60.742461	63.371686	62.685087	62.420615	60.350199	37.451996	62.440699	60.089372	
40	36.958001	61.467895	63.162824	66.068580	65.600754	65.052675	63.045537	36.574081	65.356386	62.916692	
45	36.236114	64.132421	65.755562	70.662958	69.265506	68.084082	66.258529	34.955595	68.438310	65.960888	

```
In [4]:
```

```
Out[4]:
```

Timepoint	Drug	Capomulin	Ceftamin	Infubinol	Ketapril	Naftisol	Placebo	Propriiva	Ramicane	Stelasyn	Zoniferol
0	45.000000	45.000000	45.000000	45.000000	45.000000	45.000000	45.000000	45.000000	45.000000	45.000000	45.000000
5	44.266086	46.503051	47.062001	47.389175	46.796098	47.125589	47.248967	43.944859	47.527452	46.851818	
10	43.084291	48.285125	49.403909	49.582269	48.694210	49.423329	49.101541	42.531957	49.463844	48.689881	
15	42.064317	50.094055	51.296397	52.399974	50.933018	51.359742	51.067318	41.495061	51.529409	50.779059	
20	40.716325	52.157049	53.197691	54.920935	53.644087	54.364417	53.346737	40.238325	54.067395	53.170334	

```
In [18]:
```

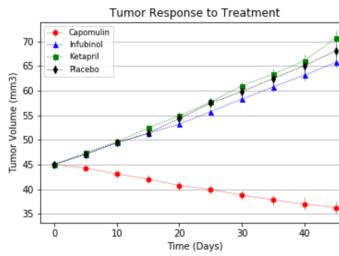
```
# Generate the Plot (with Error Bars)
drug_format_list = [('Capomulin','o','red'),('Infubinol','^','blue'),('Ketapril','s','green'),('Placebo','d','black')]
for drug,marker,color in drug_format_list:
    ste = tumor_standard_errors[drug]
    tumor_treatment_plt = plt.errorbar(pivot_avg_tumor_df.index,pivot_avg_tumor_df[drug],ste,
                                         fmt=marker,ls="--",color=color,linewidth=0.5)

plt.legend(loc ='best')
plt.title('Tumor Response to Treatment')
plt.xlabel('Time (Days)')
plt.ylabel('Tumor Volume (mm3)')
plt.grid()

# Save the Figure
plt.savefig(os.path.join('figures','tumor_response_to_treatment.png'))
```

```
File "<ipython-input-18-362f93e7cbc4>", line 8
    ^
SyntaxError: invalid syntax
```

```
In [ ]: # Show the Figure
plt.show()
```



Metastatic Response to Treatment

```
# Store the Mean Met. Site Data Grouped by Drug and Timepoint
grouped_met_df = complete_df.groupby(['Drug','Timepoint'])['Metastatic Sites'].mean()

# Convert to DataFrame
avg_met_df = grouped_met_df.reset_index()

# Preview DataFrame
avg_met_df.head()
```

```
Out[19]:
```

Drug	Timepoint	Metastatic Sites	
0	Capomulin	0	0.000000
1	Capomulin	5	0.160000
2	Capomulin	10	0.320000
3	Capomulin	15	0.375000
4	Capomulin	20	0.652174

```
In [6]:
```

```
Out[6]:
```

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

```

In [20]: # Store the Standard Error associated with Met. Sites Grouped by Drug and Timepoint
met_standard_errors = grouped.met_df.sem()

# Convert to DataFrame
met_standard_errors_df = pd.DataFrame(met_standard_errors)

# Preview DataFrame
met_standard_errors_df.reset_index(inplace=True)
met_standard_errors_df.head()

```

Out[20]:

	Drug	Timepoint	Metastatic Sites
0	Capomulin	0	0.000000
1	Capomulin	5	0.074833
2	Capomulin	10	0.125433
3	Capomulin	15	0.132048
4	Capomulin	20	0.161621


```

In [21]: # Minor Data Munging to Re-Format the Data Frames
pivot_avg_met_df = avg_met_df.pivot(index='Timepoint',columns='Drug',values='Metastatic Sites')

# Preview that Reformating worked
pivot_avg_met_df.head(20)

```

Out[21]:

	Drug	Capomulin	Ceftamin	Infubinol	Ketapril	Naftisol	Placebo	Propriva	Ramicane	Stelasyn	Zoniferol
0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
5	0.160000	0.380952	0.280000	0.304348	0.260870	0.375000	0.320000	0.120000	0.240000	0.166667	
10	0.320000	0.600000	0.666667	0.590909	0.523810	0.833333	0.565217	0.250000	0.478261	0.500000	
15	0.375000	0.789474	0.904762	0.842105	0.857143	1.250000	0.764706	0.333333	0.782609	0.809524	
20	0.652174	1.111111	1.050000	1.210526	1.150000	1.026316	1.000000	0.347826	0.952381	1.294118	
25	0.818182	1.500000	1.277778	1.631579	1.500000	1.941176	1.357143	0.652174	1.157895	1.687500	
30	1.090909	1.937500	1.588235	2.055556	2.066667	2.266667	1.615385	0.782609	1.388889	1.933333	
35	1.181818	2.071429	1.666667	2.294118	2.266667	2.642857	2.300000	0.952381	1.562500	2.285714	
40	1.380952	2.357143	2.100000	2.733333	2.466667	3.166667	2.777778	1.100000	1.583333	2.785714	
45	1.476190	2.692308	2.111111	3.363636	2.538462	3.272727	2.571429	1.250000	1.727273	3.071429	


```

In [22]: # Generate the Plot (with Error Bars)
for drug,marker,colors in drug_format_list:
    ste = met_standard_errors[drug]
    met_treatment_plt = plt.errorbar(pivot_avg_met_df.index,pivot_avg_met_df[drug],ste,
                                      fmt=marker,ls='--',color=colors,linewidth=0.5)

# Save the Figure
plt.legend(loc='best')
plt.title('Metastatic Spread During Treatment')
plt.xlabel('Treatment Duration (Days)')
plt.ylabel('Met. Sites')
plt.grid()

# Show the Figure
plt.savefig(os.path.join('figures','metastic_spread_during_treatment.png'))

No handles with labels found to put in legend.

-----
FileNotFoundError                         Traceback (most recent call last)
<ipython-input-22-4ebc5b78457b> in <module>
      13
      14 # Show the Figure
--> 15 plt.savefig(os.path.join('figures','metastic_spread_during_treatment.png'))

/opt/anaconda3/lib/python3.7/site-packages/matplotlib/pyplot.py in savefig(*args, **kwargs)
   720 def savefig(args, **kwargs):
   721     fig = gcf()
--> 722     res = fig.savefig(*args, **kwargs)
   723     fig.canvas.draw_idle() # need this if 'transparent=True' to reset colors
   724     return res

/opt/anaconda3/lib/python3.7/site-packages/matplotlib/figure.py in savefig(self, fname, transparent, **kwargs)
  2178         self.patch.set_visible(frameon)
  2179
--> 2180         self.canvas.print_figure(fname, **kwargs)
  2181
  2182         if frameon:

/opt/anaconda3/lib/python3.7/site-packages/matplotlib/backend_bases.py in print_figure(self, filename, dpi, facecolor, edgecolor, orientation, format, bbox_inches, **kwargs)
  2080             orientation=orientation,
  2081             bbox_inches_restore=bbox_inches_restore,
--> 2082             **kwargs)
  2083
  2084         finally:
  2085             if bbox_inches and restore_bbox:

/opt/anaconda3/lib/python3.7/site-packages/matplotlib/backends/backend_agg.py in print_png(self, filename_or_obj, metadata, pil_kw_args, *args, **kwargs)
  528         renderer = self.get_renderer()
  529         with cbook.setattr_cm(renderer, dpi=self.figure.dpi), \
--> 530             cbook.open_file_cm(filename_or_obj, "wb") as fh:
  531             _png.write_png(renderer, renderer, fh,
  532                           self.figure.dpi, metadata=metadata)

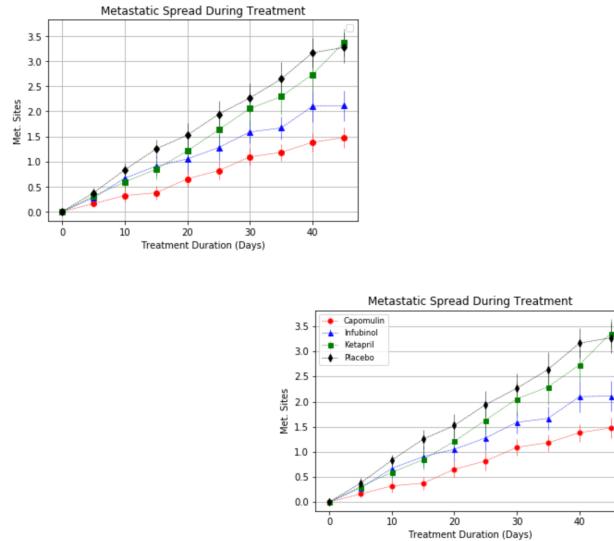
/opt/anaconda3/lib/python3.7/contextlib.py in __enter__(self)
  110     del self.args, self.kwds, self.func
  111
--> 112     try:
  113         return next(self.gen)
  114     except StopIteration:
  115         raise RuntimeError("generator didn't yield") from None

/opt/anaconda3/lib/python3.7/site-packages/matplotlib/cbook/__init__.py in open_file_cm(path_or_file, mode, encoding)
  445 def open_file_cm(path_or_file, mode='r', encoding=None):
  446     r"""Pass through file objects and context-manage `PathLike`\s."""
--> 447     fh, opened = to_filehandle(path_or_file, mode, True, encoding)
  448     if opened:
  449         with fh:

/opt/anaconda3/lib/python3.7/site-packages/matplotlib/cbook/__init__.py in to_filehandle(fname, flag, return_opened, encoding)
  430         fh = bz2.BZ2File(fname, flag)
  431     else:
--> 432         fh = open(fname, flag, encoding=encoding)
  433     opened = True
  434     elif hasattr(fname, 'seek'):

FileNotFoundError: [Errno 2] No such file or directory: 'figures/metastic_spread_during_treatment.png'

```



Survival Rates

```
In [23]: # Store the Count of Mice Grouped by Drug and Timepoint (W can pass any metric)
mouse_grouped_df = complete_df.groupby(['Drug','Timepoint'])['Mouse ID']
count_mouse_df = mouse_grouped_df.nunique()
count_mouse_df = pd.DataFrame(count_mouse_df)

# Convert to DataFrame
count_mouse_df.reset_index(inplace=True)
count_mouse_df.rename(columns={'Mouse ID':'Mouse Count'})

# Preview DataFrame
count_mouse_df.head()
```

```
Out[23]:
Drug Timepoint Mouse Count
0 Capomulin 0 25
1 Capomulin 5 25
2 Capomulin 10 25
3 Capomulin 15 24
4 Capomulin 20 23
```

```
In [24]: # Minor Data Munging to Re-Format the Data Frames
pivot_count_mouse_df = count_mouse_df.pivot(index='Timepoint',columns='Drug',values='Mouse Count')

# Preview the Data Frame
pivot_count_mouse_df.head()
```

```
Out[24]:
Drug Capomulin Cetamin Infubinol Ketapril Naftisol Placebo Propriva Ramicane Stelasyne Zoniferol
Timepoint
0 25 25 25 25 25 25 25 25 25 25 25
5 25 21 25 23 23 24 24 25 24 24 24
10 25 20 21 22 21 24 22 24 22 22 22
15 24 19 21 19 21 20 16 24 22 22 21
20 23 18 20 19 20 19 16 23 20 17
```

```
In [25]: # Generate the Plot (Accounting for percentages)
for drug,marker,colors in drug_format_list:
    total_mouse = pivot_count_mouse_df[drug][0]
    survival_rate = (pivot_count_mouse_df[drug]/total_mouse)*100
    survival_treatment_plt = plt.plot(pivot_count_mouse_df.index,survival_rate,
                                       marker=marker,ls='--',color=colors,linewidth=0.5)

# Show the Figure
plt.legend(loc='best')
plt.title('Survival During Treatment')
plt.xlabel('Times (Days)')
plt.ylabel('Survival Rate (%)')
plt.grid()

# Save the Figure
plt.savefig(os.path.join('figures','survival_during_treatment.png'))
```

No handles with labels found to put in legend.

```
----- Traceback (most recent call last)
<ipython-input-25-853e78f5c901> in <module>
    14
    15 # Save the Figure
--> 16 plt.savefig(os.path.join('figures','survival_during_treatment.png'))
/opt/anaconda3/lib/python3.7/site-packages/matplotlib/pyplot.py in savefig(*args, **kwargs)
    720     def savefig(*args, **kwargs):
    721         fig = gcf()
--> 722         res = fig.savefig(*args, **kwargs)
    723         fig.canvas.draw_idle() # need this if 'transparent=True' to reset colors
    724         return res

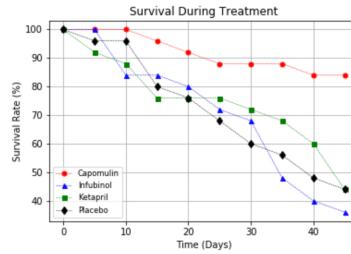
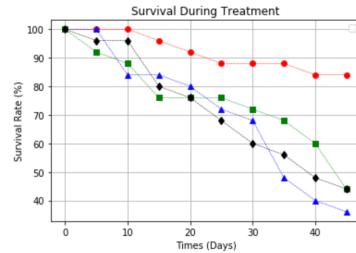
/opt/anaconda3/lib/python3.7/site-packages/matplotlib/figure.py in savefig(self, fname, transparent, **kwargs)
    2178         self.patch.set_visible(frameon)
    2179
--> 2180         self.canvas.print_figure(fname, **kwargs)
    2181
    2182         if frameon:
```

/opt/anaconda3/lib/python3.7/site-packages/matplotlib/backend_bases.py in print_figure(self, filename, dpi, facecolor, edgecolor, orientation, format, bbox_inches, **kwargs)
 2080 orientation=orientation,
 2081 bbox_inches_restore=bbox_inches_restore,
--> 2082 **kwargs)

```

2083         finally:
2084             if bbox_inches and restore_bbox:
2085                 renderer = self.get_renderer()
2086                 with cbook.setattr_cm(renderer, dpi=self.figure.dpi), \
2087                     cbook.open_file_cm(filename_or_obj, "wb") as fh:
2088                     _png.write_png(renderer._renderer, fh,
2089                                   self.figure.dpi, metadata=metadata)
2090
2091     /opt/anaconda3/lib/python3.7/contextlib.py in __enter__(self)
2092         del self.args, self.kwds, self.func
2093         try:
2094             return next(self.gen)
2095         except StopIteration:
2096             raise RuntimeError("generator didn't yield") from None
2097
2098 /opt/anaconda3/lib/python3.7/site-packages/matplotlib/cbook/_init_.py in open_file_cm(path_or_file, mode, encoding)
2099     445 def open_file_cm(path_or_file, mode='r', encoding=None):
2100         r"""Pass through file objects and context-manage .PathLike`\s."""
2101     --> 447     fh, opened = to_filehandle(path_or_file, mode, True, encoding)
2102         if opened:
2103             with fh:
2104                 fh = bz2.BZ2File(fname, flag)
2105
2106         else:
2107             fh = open(fname, flag, encoding=encoding)
2108         opened = True
2109
2110     elif hasattr(fname, 'seek'):
2111
2112 FileNotFoundError: [Errno 2] No such file or directory: 'figures/survival_during_treatment.png'

```



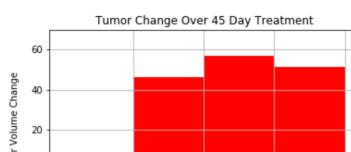
Summary Bar Graph

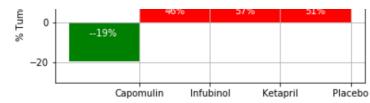
```
In [ ]: # Calculate the percent changes for each drug
# Display the data to confirm
```

```
In [13]:
```

```
Out[13]: Drug
Capomulin -19.475303
Ceftamini 42.516492
Infubinol 46.123472
Ketapril 57.028795
Naftisol 53.923347
Placebo 51.297960
Propriova 47.241175
Ramicane -22.320900
Stelasyn 52.085134
Zoniferol 46.579751
dtype: float64
```

```
In [ ]: # Store all Relevant Percent Changes into a Tuple
# Splice the data between passing and failing drugs
# Orient widths. Add labels, tick marks, etc.
# Use functions to label the percentages of changes
# Call functions to implement the function calls
# Save the Figure
# Show the Figure
fig.show()
```





In []: