**Introduction**

* **Intent of the application:** Explore the Mushroom Dataset by creating visualizations and classification trees to help identify mushrooms.
* **Dataset to be used:** Mushroom Dataset from UCI
* **Use case:**  The group of college students finally made their way to California and settled in their new housing. But because of all their money going towards rent, they need to forage for food. Exploring into the wilderness the discover a patch of wild mushrooms. Using the dataset, they find a way to create a mushroom soup that is flavorful, non-toxic, visually appealing, and pleasingly smelling.

**Dataset Analysis**

* **Define variables**:
  + cap-shape
  + cap-surface
  + cap-color
  + odor:
  + gill-attachment
  + gill-spacing
  + gill-size
  + gill-color
  + stalk-shape
  + stalk-root:
  + stalk-surface
  + stalk-surface-below-ring
  + stalk-color-above-ring
  + stalk-color-below-ring
  + veil-type
  + veil-color
  + ring-number
  + ring-type
  + spore-print-color
  + population
  + habitat
* **Define labels:**
  + cap-shape: bell=b,conical=c,convex=x,flat=f, knobbed=k,sunken=s
  + cap-surface: fibrous=f,grooves=g,scaly=y,smooth=s
  + cap-color: brown=n,buff=b,cinnamon=c,gray=g,green=r, pink=p,purple=u,red=e,white=w,yellow=y
  + bruises?: bruises=t,no=f
  + odor: almond=a,anise=l,creosote=c,fishy=y,foul=f, musty=m,none=n,pungent=p,spicy=s
  + gill-attachment: attached=a,descending=d,free=f,notched=n
  + gill-spacing: close=c,crowded=w,distant=d
  + gill-size: broad=b,narrow=n
  + gill-color: black=k,brown=n,buff=b,chocolate=h,gray=g, green=r,orange=o,pink=p,purple=u,red=e, white=w,yellow=y
  + stalk-shape: enlarging=e,tapering=t
  + stalk-root: bulbous=b,club=c,cup=u,equal=e, rhizomorphs=z,rooted=r,missing=?
  + stalk-surface-above-ring: fibrous=f,scaly=y,silky=k,smooth=s
  + stalk-surface-below-ring: fibrous=f,scaly=y,silky=k,smooth=s
  + stalk-color-above-ring: brown=n,buff=b,cinnamon=c,gray=g,orange=o, pink=p,red=e,white=w,yellow=y
  + stalk-color-below-ring: brown=n,buff=b,cinnamon=c,gray=g,orange=o, pink=p,red=e,white=w,yellow=y
  + veil-type: partial=p,universal=u
  + veil-color: brown=n,orange=o,white=w,yellow=y
  + ring-number: none=n,one=o,two=t
  + ring-type: cobwebby=c,evanescent=e,flaring=f,large=l, none=n,pendant=p,sheathing=s,zone=z
  + spore-print-color: black=k,brown=n,buff=b,chocolate=h,green=r, orange=o,purple=u,white=w,yellow=y
  + population: abundant=a,clustered=c,numerous=n, scattered=s,several=v,solitary=y
  + habitat: grasses=g,leaves=l,meadows=m,paths=p, urban=u,waste=w,woods=d

**Inputs**

* **Data import:**
  + Mushroom Data Set

**Proposed Libraries**

* **Libraries**
  + sklearn
  + Pandas
  + Pandas-profiling
  + Numpy
  + Seaborn
  + Matplotlib
* **Library source**
  + Scikit-learn.org
  + Pandas.pydata.org
  + <https://github.com/ydataai/pandas-profiling>
  + Numpy.org
  + Seaborn.pydata.org
  + https://matplotlib.org/

**Proposed Solution**

* Import the dataset and explore and display the features of the dataset.
* Explore the dataset using tools and libraries available in Python.
* Create a classification tree

**Proposed Outputs**

* Classification tree
* Confusion matrix

**Proposed Visualization**

* 3D scatterplots of mushroom color, size, etc
* Histograms

**Conclusions**

* In our analysis, we will find that certain mushrooms will have a specific smell and look that will be correlated with how toxic it is. This will be helpful in determining which stew mix the students can create. We will also see some correlation in the shapes and size of the mushrooms as well as how abundant a specific type is.