Day 2: Real-world data are not simple integers and float; Hence we need to do some pre-processing and Feature Engineering!

Topic: Feature Engineering & Pre-processing.

Resources:

1. "What Is Feature Engineering for Machine Learning?" Amit Shekhar <https://medium.com/mindorks/what-is-feature-engineering-for-machine-learning-d8ba3158d97a>  
2. "ML Crash Course by Google" Google Developers <https://developers.google.com/machine-learning/crash-course/representation/video-lecture>  
3. “Categorical Data” by Dipanjan (DJ) Sarkar <https://link.medium.com/aI7JKJSY54>  
4. “Continuous Numeric Data” by Dipanjan (DJ) Sarkar <https://link.medium.com/KAuqugRY54>  
5. “Ways to Detect and Remove the Outliers” by Natasha Sharma <https://link.medium.com/Y89PjvVY54>  
6. "Introduction to Feature Engineering" Ali Mustufa <https://colab.research.google.com/drive/1xMEqb5n1zfXJO8_lWQsovz6ubfXDmPdU#scrollTo=2TwkLcmqX-lW>

Task:

Titanic dataset cleaning+Feature Engineering and visualization only (Apply what you learned in Day1)   
Data link: <https://www.kaggle.com/c/titanic/data>

Starter pack for beginners: <https://colab.research.google.com/drive/18j97Ia-xlEKa9IWqEFt613CF1y4qgxVx>

OMKAR AGRAWAL:  
DOCS FOR REFERENCES

sklearn (Scikit-learn):  <https://scikit-learn.org/stable/getting_started.html>

keras: <https://keras.io/models/about-keras-models/>