

```
# -*- coding: utf-8 -*-
"""Cellphone User Behavior_5 Day Code
Challenge.ipynb

Automatically generated by Colab.

Original file is located at
https://colab.research.google.com/drive/1MvJZrv9TqgAajqWca9P7-0V-WLZ3AA9f
"""

!
pip install pyspan
import pandas as pd
import pyspan as
ps

pd.set_option('display.max_columns', None)
df = pd.read_csv("/content/cell
phone_user_behavior_dataset.csv")
df.head()

df.shape
df.columns

#may want to delete
columns
#drop col 'User ID'
#inspect column names that can be renamed
#clean up
'User Bhavior Class'

#remove unnecessary columns
#inspect and handle null
values

#column = 'User ID'
df = ps.remove(df, operation='columns',
columns='User ID')
df.head()

#column = 'User Behavior Class'
df =
ps.remove(df, operation='columns', columns='User Behavior
Class')
df.head()

#handle null values
pd.set_option('display.max_rows',
None)
df.isnull().sum()

#replace null values "No Loan"
df = ps.handle_nulls(df,
action = 'replace', with_val = "No Loan")

df = ps. detect_outliers(df,
method='iqr', threshold=1.5,columns='Age')
```