{"cells":[{"metadata":{},"cell_type":"markdown","source":"**This notebook is an exercise in the [Data Visualization](https://www.kaggle.com/learn/data-visualization) course. You can reference the tutorial at [this link](https://www.kaggle.com/alexisbcook/line-charts).**\n\n---\n"},{"metadata":{},"cell_type":"markdown","source":"In this exercise, you will use your new knowledge to propose a solution to a real-world scenario. To succeed, you will need to import data into Python, answer questions using the data, and generate **line charts** to understand patterns in the data.\n\n## Scenario\n\nYou have recently been hired to manage the museums in the City of Los Angeles. Your first project focuses on the four museums pictured in the images below.\n\n![ex1_museums](https://i.imgur.com/pFYL8J1.png)\n\nYou will leverage data from the Los Angeles [Data Portal](https://data.lacity.org/) that tracks monthly visitors to each museum. \n\n![ex1_xlsx](https://i.imgur.com/mGWYlym.png)\n\n## Setup\n\nRun the next cell to import and configure the Python libraries that you need to complete the exercise."},{"metadata":{"trusted":false},"cell_type":"code","source":"import pandas as pd\npd.plotting.register_matplotlib_converters()\nimport matplotlib.pyplot as plt\n%matplotlib inline\nimport seaborn as sns\nprint(\"Setup Complete\")","execution_count":null,"outputs":[]},{"metadata":{},"cell_type":"markdown","source":"The questions below will give you feedback on your work. Run the following cell to set up the feedback system."},{"metadata":{"trusted":false},"cell_type":"code","source":"# Set up code checking\nimport os\nif not os.path.exists(\"../input/museum_visitors.csv\"):\n   os.symlink(\"../input/data-for-datavis/museum_visitors.csv\", \"../input/museum_visitors.csv\") \nfrom learntools.core import binder\nbinder.bind(globals())\nfrom learntools.data_viz_to_coder.ex2 import *\nprint(\"Setup Complete\")","execution_count":null,"outputs":[]},{"metadata":{},"cell_type":"markdown","source":"## Step 1: Load the data\n\nYour first assignment is to read the LA Museum Visitors data file into `museum_data`. Note that:\n- The filepath to the dataset is stored as `museum_filepath`. Please **do not** change the provided value of the filepath.\n- The name of the column to use as row labels is `\"Date\"`. (This can be seen in cell A1 when the file is opened in Excel.)\n\nTo help with this, you may find it useful to revisit some relevant code from the tutorial, which we have pasted below:\n\n```python\n# Path of the file to read\nspotify_filepath = \"../input/spotify.csv\"\n\n# Read the file into a variable spotify_data\nspotify_data = pd.read_csv(spotify_filepath, index_col=\"Date\", parse_dates=True)\n```\n\nThe code you need to write now looks very similar!"},{"metadata":{"trusted":false},"cell_type":"code","source":"# Path of the file to read\nmuseum_filepath = \"../input/museum_visitors.csv\"\n\n# Fill in the line below to read the file into a variable museum_data\nmuseum_data = pd.read_csv(museum_filepath,index_col=\"Date\",parse_dates=True)\n\n# Run the line below with no changes to check that you've loaded the data correctly\nstep_1.check()","execution_count":null,"outputs":[]},{"metadata":{"trusted":false},"cell_type":"code","source":"# Uncomment the line below to receive a hint\n#step_1.hint()\n# Uncomment the line below to see the solution\n#step_1.solution()","execution_count":null,"outputs":[]},{"metadata":{},"cell_type":"markdown","source":"## Step 2: Review the data\n\nUse a Python command to print the last 5 rows of the data."},{"metadata":{"trusted":false},"cell_type":"code","source":"# Print the last five rows of the data\nmuseum_data.tail()","execution_count":null,"outputs":[]},{"metadata":{},"cell_type":"markdown","source":"The last row (for `2018-11-01`) tracks the number of visitors to each museum in November 2018, the next-to-last row (for `2018-10-01`) tracks the number of visitors to each museum in October 2018, _and so on_.\n\nUse the last 5 rows of the data to answer the questions below."},{"metadata":{"trusted":false},"cell_type":"code","source":"# Fill in the line below: How many visitors did the Chinese American Museum \n# receive in July 2018?\nca_museum_jul18 = 2620\n\n# Fill in the line below: In October 2018, how many more visitors did Avila \n# Adobe receive than the Firehouse Museum?\navila_oct18 = 19280-4622\n\n# Check your answers\nstep_2.check()","execution_count":null,"outputs":[]},{"metadata":{"trusted":false},"cell_type":"code","source":"# Lines below will give you a hint or solution code\n#step_2.hint()\n#step_2.solution()","execution_count":null,"outputs":[]},{"metadata":{},"cell_type":"markdown","source":"## Step 3: Convince the museum board \n\nThe Firehouse Museum claims they ran an event in 2014 that brought an incredible number of visitors, and that they should get extra budget to run a similar event again. The other museums think these types of events aren't that important, and budgets should be split purely based on recent visitors on an average day. \n\nTo show the museum board how the event compared to regular traffic at each museum, create a line chart that shows how the number of visitors to each museum evolved over time. Your figure should have four lines (one for each museum).\n\n> **

(Optional) Note**: If you have some prior experience with plotting figures in Python, you might be familiar with the `plt.show()` command. If you decide to use this command, please place it **after** the line of code that checks your answer (in this case, place it after `step_3.check()` below) -- otherwise, the checking code will return an error!"},{"metadata":
{"trusted":false},"cell_type":"code","source":"# Line chart showing the number of visitors to each museum over time\n# Line chart showing the number of visitors to each museum over time\nsns.lineplot(data=museum_data)# Your code here\n\n\n# Check your answer\nstep_3.check()","execution_count":null,"outputs":[]},{"metadata":{"trusted":false},"cell_type":"code","source":"# Lines below will give you a hint or solution code\n#step_3.hint()\n#step_3.solution_plot()","execution_count":null,"outputs":[]},
{"metadata":{},"cell_type":"markdown","source":"## Step 4: Assess seasonality\n\nWhen meeting with the employees at Avila Adobe, you hear that one major pain point is that the number of museum visitors varies greatly with the seasons, with low seasons (when the employees are perfectly staffed and happy) and also high seasons (when the employees are understaffed and stressed). You realize that if you can predict these high and low seasons, you can plan ahead to hire some additional seasonal employees to help out with the extra work.\n\n#### Part A\nCreate a line chart that shows how the number of visitors to Avila Adobe has evolved over time. (_If your code returns an error, the first thing that you should check is that you've spelled the name of the column correctly! You must write the name of the column exactly as it appears in the dataset._)"},{"metadata":
{"trusted":false},"cell_type":"code","source":"# Line plot showing the number of visitors to Avila Adobe over time\nsns.lineplot(data=museum_data[\"Avila Adobe\"])# Your code here\n#check\nstep_4.a.check()","execution_count":null,"outputs":
[]},{"metadata":{"trusted":false},"cell_type":"code","source":"# Lines below will give you a hint or solution code\n#step_4.a.hint()\n#step_4.a.solution_plot()","execution_count":null,"outputs":[]},{"metadata":
{},"cell_type":"markdown","source":"#### Part B\n\nDoes Avila Adobe get more visitors:\n- in September-February (in LA, the fall and winter months), or \n- in March-August (in LA, the spring and summer)? \n\nUsing this information, when should the museum staff additional seasonal employees?"},{"metadata":
{"trusted":false},"cell_type":"code","source":"#step_4.b.hint()","execution_count":null,"outputs":[]},{"metadata":
{"trusted":false},"cell_type":"code","source":"# Check your answer (Run this code cell to receive credit!)\nstep_4.b.solution()","execution_count":null,"outputs":[]},{"metadata":{},"cell_type":"markdown","source":"# Keep going\n\nMove on to learn about **[bar charts and heatmaps](https://www.kaggle.com/alexisbcook/bar-charts-and-heatmaps)** with a new dataset!"},{"metadata":{},"cell_type":"markdown","source":"---\n\n\n\n\n*Have questions or comments? Visit the [Learn Discussion forum](https://www.kaggle.com/learn-forum/161291) to chat with other Learners.*"}],"metadata":{"kernelspec":
{"language":"python","display_name":"Python 3","name":"python3"},"language_info":
{"pygments_lexer":"ipython3","nbconvert_exporter":"python","version":"3.6.4","file_extension":".py","codemirror_mode":
{"name":"ipython","version":3},"name":"python","mimetype":"text/x-python"}},"nbformat":4,"nbformat_minor":4}