12.010 Problem Set 2

Question 1: Develop a Graphical User Interface to download, display, and analyze GPS time series positions. There are 3 main sources of GPS time series data that show how daily position estimates of GPS stations change with time: University of Nevada (UNR), Jet Propulsion Laboratory (JPL) and Earthscope (also known as Unavco). Each group makes GPS timeseries available but they use different formats for the files (see examples below).
Hint: Look at Lec04_DataFrames.ipynb for cells that read and plotted mean sea-level data. For files that use spaces between the columns, the delimiter=r'\s+' can be used instead of delimiter=',' . Each of the files contains position differences in North, East, and Height from an arbitrary location with standard deviations of the estimates. Some use meters and other millimeters for the units. Your plots should be in millimeters.

The main features of the GUI should be

(a) Select a 4-character station code (e.g., P179) and download the data from University of Nevada (url: http://geodesy.unr.edu/gps_timeseries/tenv3/plates/NA/P179.NA.tenv3). Or from JPL
url: https://sideshow.jpl.nasa.gov/pub/JPL_GPS_Timeseries/repro2018a/post/point/P179.series

Example formats are shown below. P179 should be replaced with 4-char code from the user.
For an extra challenge, you can try downloading from UNAVCO (now the Earthscope consortium), but the National Science Foundation requires users to be authenticated, and this requires registering with UNAVCO. (see https://www.unavco.org/news/new-user-identity-management-system-for-data-access/) You can use the CiLogon option, which will allow you to use your MIT ID to log in. In a web browser, once you have connected this way, you should be able to download files through the browser. For automated processing, where your notebook can download data, you will need to follow one of the examples given here: https://www.unavco.org/data/gps-gnss/file-server/file-server-access-examples.html. Test your access with the URL below. Comments about trying to download data from a PC will be very useful feedback for the Earthscope facility.
(url: https://data.unavco.org/archive/gnss/products/position/P179/P179.cwu.nam14.csv)
Example formats are shown below. Note that the file contents are different.
Your GUI should be able to download from UNR and JPL and have the option of UNAVCO even if you can't get it to work. (A Pull-down menu would allow a service selection if using ipywidgets).

(b) Your GUI should save a local copy of the file so that once accessed; it is not necessary to re-access the web servers at JPL, UNR and UNAVCO. (The option to save a copy could be a button in the GUI). For UNAVCO, this feature would allow you to download data through a web browser and then access the file through your GUI interface.

(c) Data (in mm) should be plotted with errors with the option in the interface of selecting that only every Nth error bar is plotted. (User specifies N somewhere in the interface). Plots can be in 3 subplots or have a GUI option to select either North, East, or Up/Vertical

to be plotted. Ideally, the GUI should be able to overlay plots from different analysis centers and/or stations (especially after removing trends from part (e).

(d) Removal of data with large error bars (e.g., > ±10 mm on North and East and >30 mm in Height). The specific numerical values should be selectable in the GUI.

(e) Remove a linear fit from the data accounting for the error bars and show the estimated rate and standard deviation on the plot of residuals to the fit. Other measures of the fit to the data, e.g., (sum of residuals/sigma)^2/(numdata-2) – Chi**2 degree of freedoms and the number of data can be shown). (See Lec12_Regression.ipynb for example of how to this).

An updated version of Lec10_ipywidgets.ipynb has a GUI example for plotting equations. There is also code you can use to access Earthscope data.

Each part of the question is 20-points.

Example files:

Notice that in these files, dates are represented differently, and you will need to decide on how you want to represent dates in your GUI and plots. The smallest interval between dates is 1 day.

UNR format is: Columns labeled north(m), east(m), and up(m) should be plotted against the date. The error bars are sig_n, sig_e and sig_u. (Notice units here are meters).
http://geodesy.unr.edu/gps_timeseries/tenv3/plates/NA/P179.NA.tenv3

```
site YYMMMDD yyyy.yyyy __MJD week d reflon _e0(m) __east(m) ____n0(m) _north(m) u0(m) ____up(m)
_ant(m) sig_e(m) sig_n(m) sig_u(m) __corr_en __corr_eu __corr_nu
P179 07JUN12 2007.4442 54263 1431 2 -123.7   1193  0.672128   4662630  0.195584   402  0.204377
0.0083 0.000727 0.000945 0.002945 -0.000609 -0.022918 -0.133995
P179 07JUN13 2007.4470 54264 1431 3 -123.7   1193  0.671281   4662630  0.198351   402  0.201169
0.0083 0.000671 0.000891 0.002770  0.028067 -0.036883 -0.182603
```

... file continues for > 5000 lines. Lines are wrapped above.

JPL format is:
https://sideshow.jpl.nasa.gov/pub/JPL_GPS_Timeseries/repro2018a/post/point/P179.series

```
2007.44454902         0.235945         0.054218        -0.003356         0.000688
0.000887         0.002751         0.014140         0.000136        -0.140127     234932100.00
2007  6 12 14 55  0
2007.44694939         0.235512         0.057349        -0.005257         0.000604
0.000813         0.002477         0.033666        -0.007983        -0.181316     235007850.00
2007  6 13 11 57 30
2007.44968724         0.236892         0.056248        -0.004430         0.000585
0.000790         0.002356         0.017612        -0.025714        -0.188102     235094250.00
2007  6 14 11 57 30
2007.45242509         0.237758         0.057692        -0.008304         0.000594
0.000797         0.002398         0.033342        -0.027233        -0.178331     235180650.00
2007  6 15 11 57 30
```

… files continues for >5000 lines.

There are no headers on these files (lines above are wrapped). The columns are
Year dEast dNorth dHeight sig_e sig_n sig_u corr_en corr_eu corr_nu Seconds YR Mon Day Hr Min
Where dEast dNorth dHeight are differences in North, East and Height (meters) from a reference location, sig_n, sig_e and sig_u are standard deviations of the position estimates (meters) and

corr_en corr_eu corr_nu are correlations between the position estimates (you can ignore these correlations). The web site https://sideshow.jpl.nasa.gov/post/series.html shows all the JPL sites and a plot of the above times can be seen at https://sideshow.jpl.nasa.gov/post/links/P179.html

For UNAVCO: the North (mm), East (mm), Vertical (mm) should plotted against the date with the error bars from North Std. Deviation (mm), East Std. Deviation (mm), Vertical Std. Deviation (mm).

https://data.unavco.org/archive/gnss/products/position/P179/P179.cwu.nam14.csv

```
PBO Station Position Time Series.
Format Version, 1.2.0
Reference Frame, NAM14
4-character ID, P179
Station name, IllinoisApOR2007
Begin Date, 2007-06-13
End Date, 2024-10-12
Release Date, 2024-10-13
Source file, P179.cwu.nam14.pos
Offset from source file, 47.82 mm North, -1.36 mm East, 2.98 mm Vertical
Reference position, 42.0989719257 North Latitude, -123.6855716717 East Longitude, 402.20991
meters elevation
Date, North (mm), East (mm), Vertical (mm), North Std. Deviation (mm), East Std. Deviation (mm),
Vertical Std. Deviation (mm), Quality,
2007-06-13,0.00, 0.00, 0.00, 2.09, 1.55, 7.03, repro,
2007-06-14,0.96, 0.40, 0.30, 2.13, 1.59, 7.14, repro,
2007-06-15,2.02, 2.14, 1.18, 2.08, 1.56, 7.06, repro,
2007-06-16,0.34, 1.13, -1.18, 2.07, 1.56, 7.03, repro,
2007-06-17,1.02, 1.62, -0.61, 2.08, 1.55, 7.02, repro,
2007-06-18,0.74, 2.51, 8.92, 2.09, 1.54, 7, repro,
2007-06-19,-0.45, 1.71, 1.97, 2.14, 1.6, 7.21, repro,...
```

12.010 Computational Methods of Scientific Programming

Fall 2024