

Effect of Age and APOE on Hippocampal Subfields in Cognitively Normal Subjects (OHBM 2019)

Collect data from XNAT

```
import pyxnat
c = pyxnat.Interface(config='/home/grg/.xnat_bsc.cfg')

# Collect experiments
project = 'ALFA_OPCIONAL'
experiments = c.array.experiments(project_id=project, columns=['subject_label'])

table = []
# Querying FreeSurfer resource -if any- from each experiment

for e in experiments.data:
    s = int(e['subject_label'])
    r = c.select.experiment(e['ID']).resource('FREESURFER6')

    if r.exists():
        volumes = r.hippoSfVolumes()
        volumes['subject'] = s
        table.append(volumes)

# Convert to dataframe
import pandas as pd
hippoSfVolumes = pd.concat(table).set_index('subject').sort_index()
hippoSfVolumes.head()
```

	side	region	value
subject			
10011	left	parasubiculum	42.280882
10011	left	HATA	42.460512
10011	left	Whole_hippocampus	2624.451049
10011	left	CA3	134.967436
10011	left	GC-ML-DG	214.695151

Plot each subfield in relation to age per each APOE group

```
from roistats import plotting
from matplotlib import pyplot as plt
data = hippoSfVolumes

# replacing troublesome characters in regions
data = data.replace({'hippocampal-fissure': 'hippocampal_fissure',
                    'GC-ML-DG': 'GC_ML_DG'})

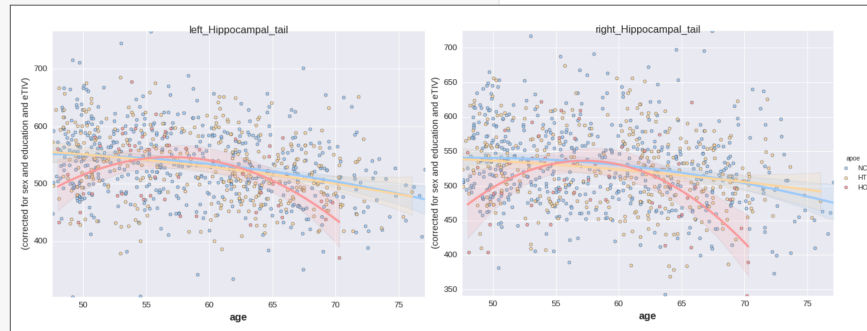
regions = [str(e) for e in set(data['region'])] # collect region names
covariates = pd.read_excel('./data/covariates.xls', index_col=0)

for region in regions[9:10]: # hippocampal tail only for this example
    for side in ['left', 'right']:
        # Selecting data from the big table
        region_with_side = '%s_%s'%(side, region)

        d = data.query('side == "%s" & region == "%s"'%(side, region))
        d = plotting.pivot(d, covariates=[]).join(covariates).dropna()
        d = d.rename(columns={region : region_with_side})

        _ = plotting.lmplot(region_with_side, 'age', d,
                           covariates=['sex', 'education', 'eTIV'],
                           hue='apoe', order = 2, palette='apoe')

        fig = plt.figure()
```

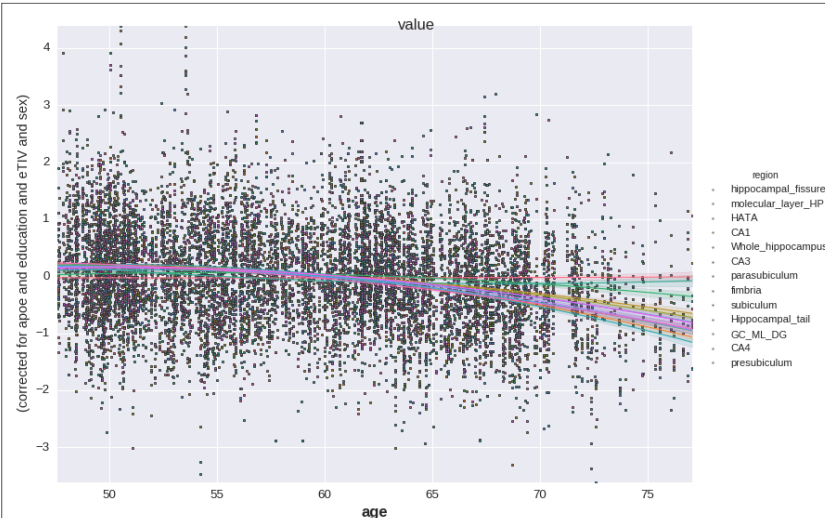


Plot each subfield in all subjects in relation to age

```
d = data.query('side == "%s" %side').join(covariates).dropna()

# converting to z-scores
dz = []
for region in regions:
    region_data = d.query('region == "%s" %region')[['region', 'value']]
    m, s = region_data['value'].mean(), region_data['value'].std()
    d2 = pd.DataFrame(region_data, columns = ['region', 'value'])
    d2['value'] = (d2['value'] - m )/s
    dz.append(d2)
dz = pd.concat(dz)
del d['region']
del d['value']
d = d.join(dz)

_ = plotting.lmplot('value', 'age', d,
                   covariates=['apoe', 'education', 'eTIV', 'sex'],
                   hue='region', order=2, s=10, size=20)
```



Collaborative quality control performed with

snaprate

github.com/xgrg/snaprate
