Exploring defining single value indices - SPI

This notebook contains several explorations and developments leading to the SPI framework.

Setup

Import Libraries

```
import soundscapy as sspy
import matplotlib.pyplot as plt
import pandas as pd
from pathlib import Path
import seaborn as sns
import utils
import rpyskewnorm as snpy
import numpy as np
from MultiSkewNorm import MultiSkewNorm

import warnings

warnings.filterwarnings("ignore")
```

Load Data

In addition to loading the latest version of the ISD, we also exclude a few samples that were identified as survey outliers. Most notably, this includes the samples at RegentsParkFields which were impacted by helicopter flyovers.

```
# Load latest ISD dataset

data = sspy.isd.load()
data, excl_data = sspy.isd.validate(data)
data = data.query("Language != 'cmn'")

# Exclude RegentsParkJapan outliers
# excl_id = list(data.query("LocationID == 'RegentsParkJapan'").query("ISOEventful > 0.72 | I
# Excluded RegentsParkFields outliers
# excl_id = excl_id + list(data.query("LocationID == 'RegentsParkFields' and ISOPleasant < 0
excl_id = [652, 706, 548, 550, 551, 553, 569, 580, 609, 618, 623, 636, 643]
data.drop(excl_id, inplace=True)
data
```

```
Renaming PAQ columns.
Checking PAQ data quality.
Identified 109 samples to remove.
[6, 9, 13, 30, 32, 46, 190, 213, 229, 244, 296, 412, 413, 428, 464, 485, 655, 734, 739, 762,
```

	LocationID	SessionID	GroupID	RecordID	start_time	end_time	latitude	longitude	Language
0	CarloV	CarloV2	2CV12	1434	2019-05-16 18:46:00	2019-05-16 18:56			
1	CarloV	CarloV2	2CV12	1435	2019-05-16 18:46:00	2019-05-16 18:56			
2	CarloV	CarloV2	2CV13	1430	2019-05-16 19:02:00	2019-05-16 19:12			
3	CarloV	CarloV2	2CV13	1431	2019-05-16 19:02:00	2019-05-16 19:12			
4	CarloV	CarloV2	2CV13	1432	2019-05-16 19:02:00	2019-05-16 19:12			
...	...	...	...	...	...	...			
1693	Noorderplantsoen	Noorderplantsoen1	NP161	61	2020-03-11 12:42:00	2020-03-11 12:55			
1694	Noorderplantsoen	Noorderplantsoen1	NP162	63	2020-03-11 12:39:00	2020-03-11 13:00			
1695	Noorderplantsoen	Noorderplantsoen1	NP162	62	2020-03-11 12:54:00	2020-03-11 12:58			
1696	Noorderplantsoen	Noorderplantsoen1	NP162	64	2020-03-11 12:56:00	2020-03-11 12:59			
1697	Noorderplantsoen	Noorderplantsoen1	NP163	70	2020-03-11 23:08:00	2020-03-11 23:18			

ISOCoordinate calculation according to Aletta et. al. (2024)

To move the 8-item PAQ responses into the 2-dimensional circumplex space, we use the projection method first presented in ISO 12913-3:2018. This projection method and its associated formulae were recently updated further in @Aletta2024 to include a correction for the language in which the survey was conducted. The formulae are as follows:

$$P_{ISO} = \frac{1}{\lambda_{pl}} \sum_{i=1}^{8} \cos\theta_i \cdot \sigma_i E_{ISO} = \frac{1}{\lambda_{pl}} \sum_{i=1}^{8} \sin\theta_i \cdot \sigma_i$$

where PAQ_i is the response to the (i)th item of the PAQ. The resulting (x) and (y) values are then used to calculate the polar angle ( ) and the radial distance (r) as follows:

```
from soundscapy.utils.parameters import LANGUAGE_ANGLES, PAQ_IDS

LANGUAGE_ANGLES
```

```
{'eng': (0, 46, 94, 138, 177, 241, 275, 340),
 'arb': (0, 36, 45, 135, 167, 201, 242, 308),
 'cmn': (0, 18, 38, 154, 171, 196, 217, 318),
 'hrv': (0, 84, 93, 160, 173, 243, 273, 354),
 'nld': (0, 43, 111, 125, 174, 257, 307, 341),
 'deu': (0, 64, 97, 132, 182, 254, 282, 336),
 'ell': (0, 72, 86, 133, 161, 233, 267, 328),
 'ind': (0, 53, 104, 123, 139, 202, 284, 308),
 'ita': (0, 57, 104, 143, 170, 274, 285, 336),
 'spa': (0, 41, 103, 147, 174, 238, 279, 332),
 'swe': (0, 66, 87, 146, 175, 249, 275, 335),
 'tur': (0, 55, 97, 106, 157, 254, 289, 313)}
```

```
tab = pd.DataFrame.from_dict(LANGUAGE_ANGLES, orient='index', columns=PAQ_IDS)
tab
```

Table 2: Language-specific angles for projection into the ISO 12913-3:2018 circumplex space.

	PAQ1	PAQ2	PAQ3	PAQ4	PAQ5	PAQ6	PAQ7	PAQ8
eng	0	46	94	138	177	241	275	340
arb	0	36	45	135	167	201	242	308
cmn	0	18	38	154	171	196	217	318
hrv	0	84	93	160	173	243	273	354
nld	0	43	111	125	174	257	307	341
deu	0	64	97	132	182	254	282	336
ell	0	72	86	133	161	233	267	328
ind	0	53	104	123	139	202	284	308
ita	0	57	104	143	170	274	285	336
spa	0	41	103	147	174	238	279	332
swe	0	66	87	146	175	249	275	335

Table 2: Language-specific angles for projection into the ISO 12913-3:2018 circumplex space.

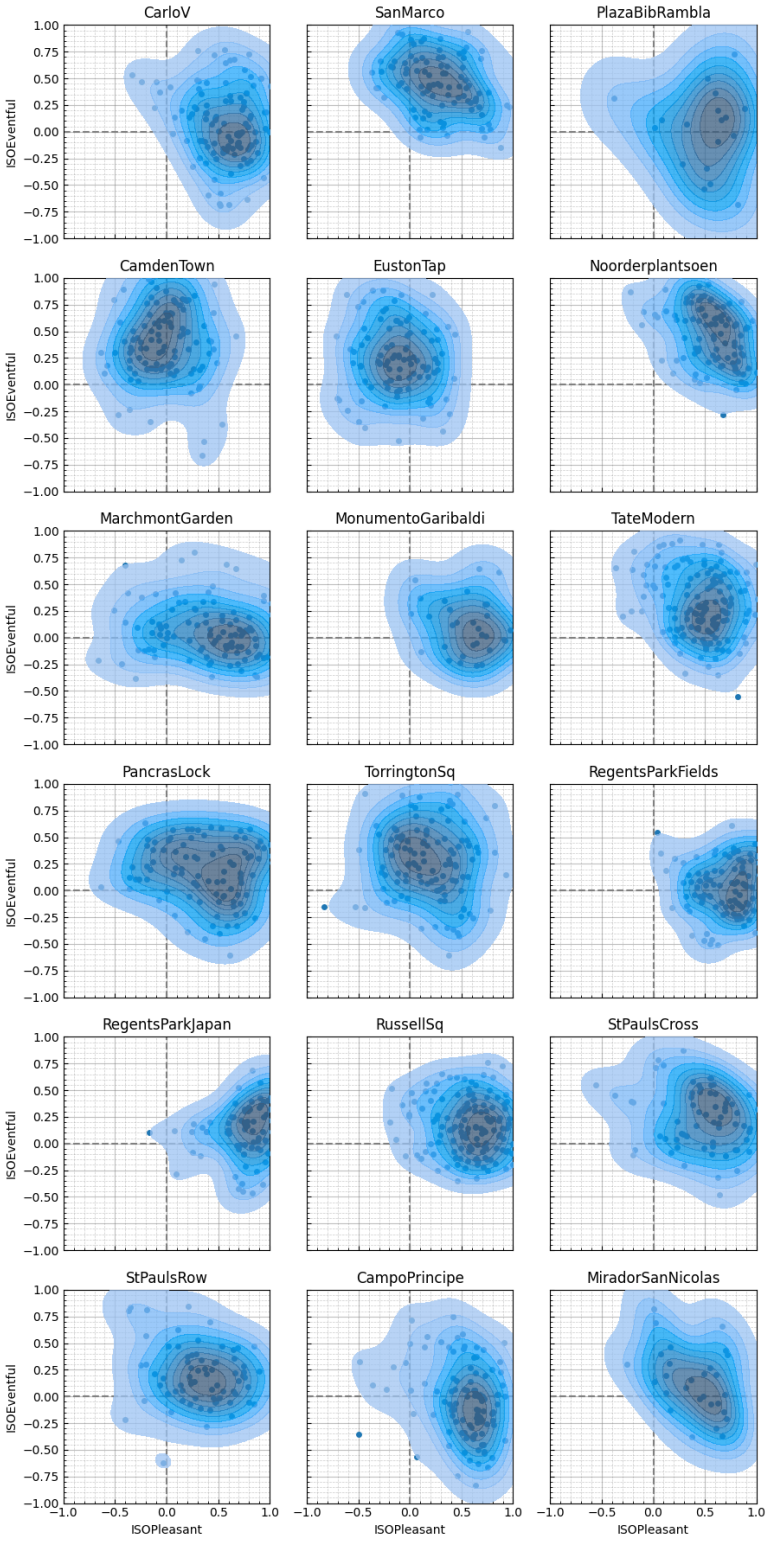
	PAQ1	PAQ2	PAQ3	PAQ4	PAQ5	PAQ6	PAQ7	PAQ8
tur	0	55	97	106	157	254	289	313

```
from soundscapy.utils.parameters import PAQ_IDS

for i, row in data.iterrows():
    lang = row["Language"]
    angles = LANGUAGE_ANGLES[lang]
    iso_pl, iso_ev = (
        sspy.surveys.adj_iso_pl(row[PAQ_IDS], angles, scale=4),
        sspy.surveys.adj_iso_ev(row[PAQ_IDS], angles, scale=4),
    )
    data.loc[i, "ISOPleasant"] = iso_pl
    data.loc[i, "ISOEventful"] = iso_ev

fig, axes = plt.subplots(6, 3, figsize=(9, 18), sharex=True, sharey=True)
for i, (loc, ax) in enumerate(zip(data.LocationID.unique(), axes.flatten())):
    sspy.plotting.density(
        data.query(f"LocationID == '{loc}'"),
        ax=ax,
        title=loc,
        # hue='Language',
    )

fig.tight_layout()
```


CarloV
SanMarco
PlazaBibRambla
CamdenTown
EustonTap
Noorderplantsoen
MarchmontGarden
MonumentoGaribaldi
TateModern
PancrasLock
TorringtonSq
RegentsParkFields
RegentsParkJapan
RussellSq
StPaulsCross
StPaulsRow
CampoPrincipe
MiradorSanNicolas
ISOEventful
ISOPleasant
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
−1.0
−0.5
0.0
0.5
1.0