

INVESTIGATION

DATA SHAPING WITH PYTHON FOR TABLEAU

Kelly Jones | Metis | April 17, 2018

CHALLENGE

Python's powers of
data manipulation



Tableau's powers of
data visualization



How can we take advantage of both?

RECOMMENDATIONS

- In Python, intentionally shape data for Tableau
 1. Long & skinny, not short & wide
 2. Do not pre-aggregate
- Extend functionality of Tableau with TabPy

✗ **AVOID THIS** ✗

Student	Project 1	Project 2	Project 3	Project 4	Final Grade
Agatha	80	84	90	96	87.5
Bertrand	70	95	90	85	85
Charlene	65	75	95	85	80
Dominic	50	55	75	80	65
Ethel	60	80	100	90	82.5
CLASS AVERAGE	65	77.8	90	87.2	77.6

DO THIS INSTEAD

Student	Project	Score
Agatha	project1	80
Agatha	project2	84
Agatha	project3	90
Agatha	project4	96
Bertrand	project1	70
Bertrand	project2	95
Bertrand	project3	90
Bertrand	project4	85
Charlene	project1	65
Charlene	project2	75
Charlene	project3	95
Charlene	project4	85
Dominic	project1	50
Dominic	project2	55
Dominic	project3	75
Dominic	project4	80
Ethel	project1	60
Ethel	project2	80
Ethel	project3	100
Ethel	project4	90

DATA SHAPE

- Long & skinny means we have more useful categories
 - “Project” instead of multiple individual projects
 - “Year” instead of multiple individual years
- Tableau prefers record-level data, not aggregate

Let's explore
in Tableau

TABPY EXTENSION

- TabPy available on GitHub
- Enables complex calculations within Tableau using Python scripts

EXAMPLE: CRIME DATA CLUSTERING



Seattle_Police_Department_911_Incident_Response ×

Results are computed along At Scene Time.

```
SCRIPT_STR("import numpy as np
from sklearn.cluster import DBSCAN
X=np.column_stack([np.radians(_arg1),np.radians(_arg2)])
db = DBSCAN(eps=_arg3[1], min_samples=_arg4[1], metric='haversine').fit(X)
return np.where(db.labels_ == np.array(-1), 'No', 'Yes').tolist()",
AVG([Latitude]),AVG([Longitude]),
[Distance between incidents],AVG([Incident Count]))
```

The calculation is valid.

Apply

OK

Source: Bora Beran, “Building advanced analytics applications with TabPy” (available [here](#))

Questions?

Thank you!