

## DataEng S24: Data Integration In-class Assignment

### A. Discussion Question

Develop an integrated alert system that combines a person's health history, such as allergies or chronic illnesses, with seasonal and weather information. This system would analyze when certain seasons or weather patterns are expected to occur and send notifications to the person, reminding them to take necessary precautions to manage their health conditions effectively.

### B. Transform the ACS Data

County	Population	%poverty	PerCapitaIncome
Loudoun County, Virginia	374558	3.689598	50455.645745
Washington County, Oregon	572071	10.321202	35369.047499
Harlan County, Kentucky	27548	35.66948	15456.971032
Malheur County, Oregon	30421	24.298225	17567.504323

Answer the following questions:

- Most populous county in the USA?  
Most populous county in the USA:ID  
State California  
County Los Angeles  
Total Population 10105722  
Poverty % 16.998271  
Per Capita Income 30798.189407  
Name: 782, dtype: object

783

- Least populous county in the USA?

```
Least populous county in the USA:ID      2878
State      Texas
County      Loving
Total Population      74
Poverty %      17.1
Per Capita Income      35530.0
Name: 2877, dtype: object
```

### C. Transform the COVID Data

County	Month	# cases	# deaths
Malheur County, Oregon	August 2020	28163	459.0
Malheur County, Oregon	January 2021	96297	1627.0
Malheur County, Oregon	February 2021	65951	1137.0

#### D. Integrate COVID Data with ACS Data

County	Poverty %	TotalCasesPer100K
Washington County, Oregon	10.321202	377110.358679
Malheur County, Oregon	24.298225	1.491187e+06
Loudoun County, Virginia	3.689598	666505.58792
Harlan County, Kentucky	35.669482	747727.60273

#### E. Analysis

County	R value
For Oregon Counties only: correlation between % poverty and COVID cases	0.28707860802137736
For all counties: correlation between population and COVID cases	0.9662049145438587
For Oregon counties only: correlation between PerCapitaIncome and COVID deaths	-0.46186659505185557
For all USA counties: correlation between PerCapitaIncome and COVID cases	-0.21366063360627474
For all USA counties: correlation between Total cases and Total Deaths	0.9044409109065387
For all Oregon counties: correlation between Total cases and Total Deaths	0.97479855471827513

## F. Charting

