IPD things

CVs

1. The Female indicator is the only indicator with acceptable CVs. One may be able to advocate for Older Adults and Youth if they're willing to tolerate CVs up to 30%. (Table 1)
2. The threshold for an acceptable CV is somewhere between 10-12%. Only the Female, Older Adults, and Youth indicators hit this mark regarding their median CVs. (Table 2)
3. The Ethnic Minority and LEP indicators are not reliable. More than 20% of tracts in the EM and LEP datasets have CVs over 60%. (Table 1)
4. Using unreliable data to make decisions has real consequences for our work. Table 3 shows the 25th, 50th, 75th, and 90th percentile LEP observations for our region.

   The only place where the possible range of LEP speakers doesn't overlap is the difference between the 75th and 90th percentile observations. I suspect DVRPC is federally mandated to provide services not only to neighborhoods with incredibly high (90th+ percentile) estimated counts of LEP speakers, but also to neighborhoods with above-average (75th percentile) estimated counts of speakers. Table 3 shows that we have significant overlap between the possible values of the 50th and 75th percentile estimates.

   When we base our decisions on estimates with such wide spread, how can we know we're providing translation services where they're actually needed? Whose needs are we neglecting?

   Remember that this is the overall LEP dataset. I would be *floored* if the language-specific data were *more* reliable than the LEP dataset referenced in the above table. And again, the LEP dataset is *not* reliable.

5. Census tracts with higher IPD scores tend to have *worse* data reliability. Of the 115 census tracts with IPD scores of 25 or higher, roughly one-third have CVs above 15%. (Table 4)

Table 1

	Indicator								
CV	D	EM	F	FB	LEP	LI	OA	RM	Y
0-15%	298 (21.78%)	90 (6.58%)	1360 (99.42%)	145 (10.6%)	57 (4.17%)	369 (26.97%)	1060 (77.49%)	615 (44.96%)	1045 (76.39%)
15.1-30%	1024 (74.85%)	230 (16.81%)	5 (0.37%)	657 (48.03%)	309 (22.59%)	817 (59.72%)	281 (20.54%)	445 (32.53%)	293 (21.42%)
30.1-60%	41 (3%)	769 (56.21%)	1 (0.07%)	515 (37.65%)	716 (52.34%)	177 (12.94%)	20 (1.46%)	290 (21.2%)	22 (1.61%)
60.1+%	5 (0.37%)	279 (20.39%)	2 (0.15%)	51 (3.73%)	286 (20.91%)	5 (0.37%)	7 (0.51%)	18 (1.32%)	8 (0.58%)

Table 2

	Indicator								
CV	D	EM	F	FB	LEP	LI	OA	RM	Y
Min	8.77	4.26	2.28	2.00	7.47	3.02	3.33	0.06	1.86
Med	17.44	42.51	4.57	27.43	40.53	19.65	10.66	16.42	10.16
Mean	18.70	44.93	5.06	29.90	44.75	20.71	12.38	20.17	12.11
Max	104.64	182.37	97.26	121.58	151.98	144.85	100.00	121.58	138.95

Table 3

Percentile	LEP Pct. Estimate	Lower Bound	Upper Bound
25	1.8% (± 1.54%)	0.26%	3.34%
50	3.6% (± 2.61%)	0.99%	6.21%
75	7.0% (± 4.16%)	2.84%	11.16%
90	16.4% (± 4.0%)	12.40%	20.40%

Table 4

	IPD Score			
Mean Reliability	0-12	13-18	19-24	25-36
0-15%	79 (84.95%)	675 (86.87%)	216 (56.40%)	75 (65.22%)
15.1-30%	11 (11.83%)	87 (11.2%)	156 (40.73%)	39 (33.91%)
30.1-60%	2 (2.15%)	10 (1.29%)	9 (2.35%)	1 (0.87%)
60.1+%	1 (1.08%)	5 (0.64%)	2 (0.52%)	0 (0%)
Total	93	777	383	115

IPD Scoring Error

1. The Racial Minority indicator is the only indicator with acceptable scoring error. (Table 5)
2. Some indicators have very high scoring error. For example, the Female indicator has a mean scoring error of 39%. Another way of saying this is that there's a 39% chance a census tract that scores a 3 on the Female indicator should actually score 4 or 2. (Table 5)
3. The scoring error increases when the binning procedure is not well-calibrated to the distribution of the underlying data.
4. We get higher classification error when our class breaks draw distinctions where there aren't differences. The classification error of the Racial Minority indicator is low because this indicator has a bimodal distribution. In the case of the DVRPC region, most census tracts have very low percentages of RM residents, some tracts have high percentages of RM residents, and few are anything near a 50-50 split. (Figure 2)
5. There is no relationship between data reliability (CVs) and IPD scoring error. Therefore, it is unlikely that improving the reliability of the underlying data (i.e., by aggregating adjacent and similar census tracts) will improve the scoring error. (Figure 1)
6. We must inevitably select a single binning procedure for all indicators, and the binning procedure will fit some indicators better than others. My question is, is a 23% overall scoring error the best we can do?

Figure 1

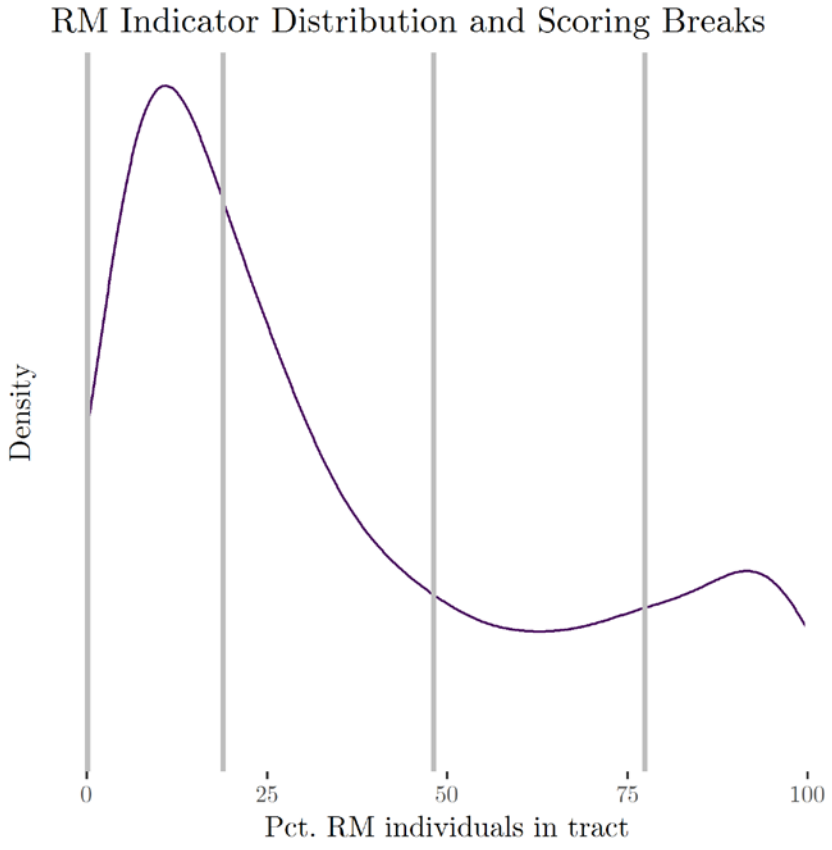

Figure 2

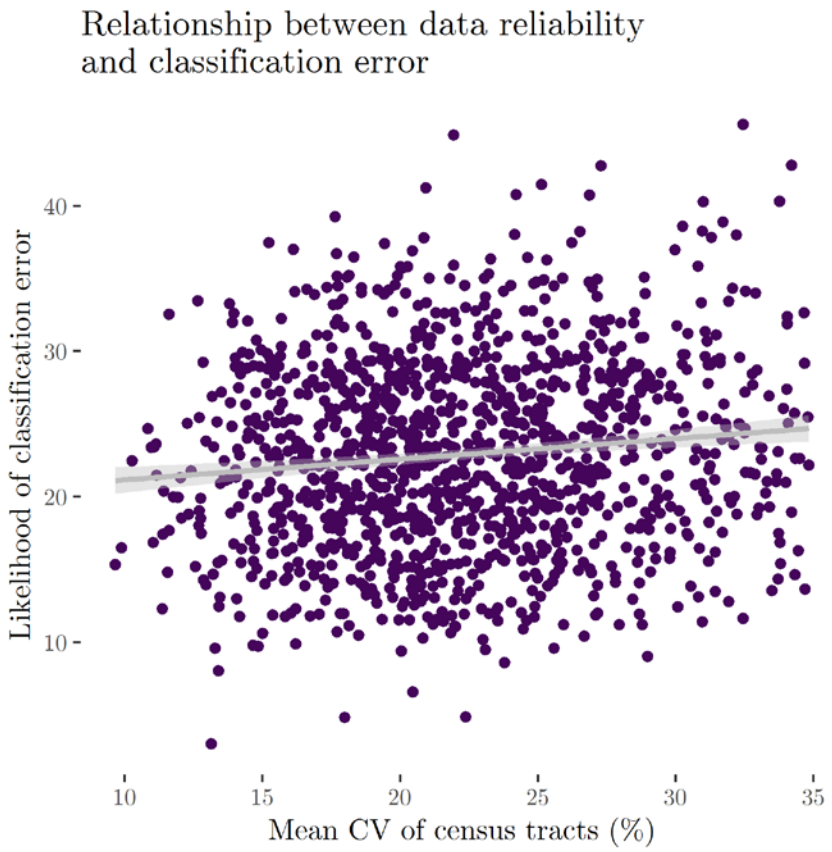

Table 5

Indicator	Mean Scoring Error
D	30.04
EM	21.663
F	39.116
FB	20.818
LEP	23.624
LI	16.244
OA	18.059
RM	9.88
Y	28.93

Improvements

It appears that improving data reliability and reducing IPD scoring error will require two separate strategies. I think it's smart to use the data-driven regionalization algorithm to create optimal geographies (aggregations of adjacent and similar census tracts) that improve the reliability of our data. This will make maps less misleading and could potentially improve the allocation of whatever resources are based upon IPD criteria, e.g. the LEP maps.

However, given what I've seen so far, I think it's unlikely that improving the data reliability will reduce IPD scoring error. Reducing IPD scoring error will require a second look at the existing binning procedure. Determining the best binning procedure among the list of common break classification schemes is theoretically simple: plug in all the indicator datasets for recent years for equal interval, quantile, natural breaks, and IPD standard deviation schemes and see what scheme is the best overall. But I know that the IPD standard deviation scheme was selected for its communicability to friends of the firm.

Prioritizing improvements should likely happen along a continuum of impact and likelihood of implementation. For example, I mention the CVs of the LEP data for two reasons. First, in terms of *impact*, I think that if we're using tract-level data to decide what neighborhoods receive translated documents and translation services, then some neighborhoods that could really use these services are being deprioritized. Second, in terms of *likelihood of implementation*, I think a better approach to our equity work regarding LEP populations is a matter of including it in a future LEP plan. Serving these populations is also a federal mandate.

A second example is changing the IPD binning procedure. In terms of *impact*, the newest draft of our TIP evaluation criteria shows the Equity criteria as 5% of the total score. I'm not sure if this is the right math (end of the day, too tired), but my back-of-the-envelope calculation shows that a given census tract is expected to be ±1.04 points off based on IPD scoring error. Is this fluctuation meaningful in terms of what projects get funded?

Quoting from a draft of the newest TIP evaluation criteria: "If project increases vehicle speeds above 35 mph or traffic volumes in above average of well above average EJ Community = 0 points. For all other projects, EJ Population Score = Σ[For all Census tracts that project is located in] Census Tract Population × IPD Score ÷ 36. Project with Highest EJ Population Score = 1 point; for all other projects: EJ Population Score ÷ Highest EJ Population Score."

Assuming that each project affects the same amount of people, the fluctuation could affect the relative order of projects that receive funding. It's also worth looking into this a little more before making a decision. However, the *impact* appears a bit more esoteric.

In terms of *likelihood of implementation*, I anticipate pushback to changing the IPD binning mechanism.