

CLASSIFYING EVICTION DATA IN OKLAHOMA

Data and code at bit.ly/ojo-at-gu

RYAN GENTZLER

Research Director

rgentzler@okpolicy.org



**OPEN
JUSTICE
OKLAHOMA**

A program of Oklahoma
Policy Institute

ABOUT OPEN JUSTICE OKLAHOMA

- Grew out of work on criminal justice issues at Oklahoma Policy Institute
- Established as program in 2018
- Maintain database of court records scraped from web
- Use data to inform our policy advocacy work, expanding to policy and program evaluation

The screenshot shows a repeating header pattern for the "OKLAHOMA State Courts Network". Below the header, a detailed court record is displayed for Case No. SC-2017-2633, a Small Claims case involving INDEBTEDNESS <\$1500. The record includes the parties (Quick Loans Inc., Plaintiff; Gary M House, Defendant), attorneys (None), events (Event: Thursday, March 2, 2017 at 9:00 AM, HEARING), and issues (Issue # 1: INDEBTEDNESS <\$1500, Filed By: Quick Loans Inc., Filed Date: 02/07/2017). A note at the bottom states: "For cases filed before 1/1/2000, ancillary issues may not appear except in the docket."

COURT RECORD STRUCTURE (1/3)

IN THE DISTRICT COURT IN AND FOR TULSA COUNTY, OKLAHOMA

Vintage Housing Inc Dba,
Plaintiff, and
Heartland Village,
Plaintiff,
v.
[REDACTED]
Defendant.

No. SC-2020-5
(Small Claims: FORCIBLE ENTRY & DETAINER <\$5000.00.)

Filed: 01/02/2020
Closed: 01/10/2020

Judge: Small Claims Judge

| | |
|--|-------------|
| | case |
| | id |
| | title |
| | district |
| | case_type |
| | year |
| | case_number |
| | date_filed |
| | date_closed |

PARTIES

[REDACTED] Defendant
Heartland Village, Plaintiff
Vintage Housing Inc Dba, Plaintiff

ATTORNEYS

Attorney
FRIERSON, N BLAINE (Bar #13979)
1126 SOUTH FRANKFORT AVE
TULSA, OK 74120

Represented Parties
Heartland Village,

| | |
|--|------------|
| | attorney |
| | id |
| | name |
| | bar_number |
| | address |
| | parties |

EVENTS

Event
Friday, January 10, 2020 at 2:00 PM
HEARING

| Party | Docket | Reporter |
|-------|---------------------|----------|
| | Deborah Ludi Leitch | |



COURT RECORD STRUCTURE (2/3)

ISSUES

For cases filed before 1/1/2000, ancillary issues may not appear except in the docket.

- Issue # 1.** Issue: FORCIBLE ENTRY & DETAINER <\$5000.00. (SCFED1)
Filed By: Aery, Warren
Filed Date: 01/02/2020

Party Name

Disposition Information

Defendant: Er [REDACTED]

Defendant: M

Disposed: DEFAULT JUDGEMENT 01/09/2020 Default Judgment

DISPOSED: DEFAULT JUDGEMENT 1.01/09/2020. Default Judgement

DOCKET

| Date | Code | Description | Party | Count | Amount |
|------------|------------|--|-------|-------|----------|
| 01-02-2020 | [TEXT] | SMALL CLAIMS INITIAL FILING. | | | #1 |
| 01-02-2020 | [SCFED1] | FORCIBLE ENTRY & DETAINER <\$5000.00. | | | |
| 01-02-2020 | [AFDC1] | AFFIDAVIT Document Available (#104570-727) TIFF PDF | | | \$ 45.00 |
| 01-02-2020 | [DMFE] | DISPUTE MEDIATION FEE | | | \$ 5.00 |
| 01-02-2020 | [PFE7] | LAW LIBRARY FEE | | | \$ 6.00 |
| 01-02-2020 | [SMIP] | SUMMONS ISSUED - PRIVATE PROCESS SERVER | | | |
| 01-02-2020 | [A/] | AMOUNT IN DEBT OF...>600.00 POSS | | | |
| 01-02-2020 | [TEXT] | OCIS HAS AUTOMATICALLY ASSIGNED SMALL CLAIMS JUDGE TO THIS CASE. | | | |

Documents contain address

Minute contains amount of debt



COURT RECORD STRUCTURE (3/3)

AFFIDAVIT

Document Available (#1045704727) [TIFF](#) [PDF](#)

01-02-2020 [DMFE]

DISPUTE MEDIATION FEE

\$ 7.00

01-02-2020 [PFE7]

LAW LIBRARY FEE

\$ 6.00

01-02-2020 [SMIP]

SUMMONS ISSUED - PRIVATE PROCESS SERVER

01-02-2020 [A/]

AMOUNT IN DEBT OF <..\$..>600.00 POSS

01-02-2020 [TEXT]

OCIS HAS AUTOMATICALLY ASSIGNED JUDGE SMALL CLAIMS JUDGE TO THIS CASE.

01-02-2020 [ACCOUNT]

RECEIPT # 2020-4039029 ON 01/02/2020.

PAYOR: WARREN AERY TOTAL AMOUNT PAID: \$ 58.00.

LINE ITEMS:

SC-2020-1: \$45.00 ON AC01 CLERK FEES.

SC-2020-1: \$6.00 ON AC23 LAW LIBRARY FEE CIVIL AND CRIMINAL.

SC-2020-1: \$7.00 ON AC64 DISPUTE MEDIATION FEES CIVIL ONLY.

01-03-2020 [FEDS]

FORCIBLE ENTRY AND DETAINER SERVED PER SERV [REDACTED] 1/2/20

Document Available (#1045705811) [TIFF](#) [PDF](#)

01-03-2020 [FEDS]

FORCIBLE ENTRY AND DETAINER SERVED MELISSA MCCELLON BY SERV [REDACTED] 1/2/20

Document Available (#1045705811) [TIFF](#) [PDF](#)

01-09-2020 [DIS]

LUDI LEITCH, DEBORRAH: PLAINTIFF APPEARED. DEFENDANTS APPEARED NOT AFTER VALID SERVICE. DEFAULT-JUDGMENT GRANTED TO THE PLAINTIFF AS FOLLOWS: (1) THE PRINCIPAL JUDGMENT AMOUNT SET FORTH HEREIN, (2) POSSESSION OF THE PREMISES AND (3) COURT COSTS AGAINST THE NAMED DEFENDANTS. JUDGE SIGNED FINAL J.E. --- PRINCIPAL JUDGMENT AMOUNT = \$600.00

#1

01-13-2020 [JES]

JOURNAL ENTRY OF JUDGMENT

Document Available (#1045702172) [TIFF](#) [PDF](#)

Minute contains
disposition detail

OUR DATA CLEANING TOOLBOX

- We do almost everything in R
- Key packages:
 - `dplyr` for categorizing
 - `tidyverse` for summarizing and grouping
 - `stringr` for string manipulation
 - `lubridate` for working with dates



CLUSTERING AND CLASSIFYING: PLAINTIFFS

Question: Which plaintiffs are filing the most evictions?

Problem: Plaintiff names are hand-entered by clerks and are non-uniform. The same plaintiff can be entered several different ways, undercounting a plaintiff's filings.

| | plaintiff | n |
|---|---|------|
| 1 | cobb | All |
| 1 | Cobblestone Apartments | 1299 |
| 2 | COBBLESTONE APARTMENTS | 507 |
| 3 | COBBLESTONE APTS | 84 |
| 4 | Cobblestone Apartments Of Tulsa, L.l.c. dba | 36 |
| 5 | Cobblestone Apartments Of Tulsa LLC DBA | 19 |
| 6 | Cobblestone Apts | 15 |
| 7 | Cobblestone Apartments Of Tulsa Llc Dba | 2 |
| 8 | Cobb, Peggy R | 1 |
| 9 | Coblesstone Apartments | 1 |

FOLLOW ALONG!
bit.ly/ojo-at-gu



CLUSTERING AND CLASSIFYING: PLAINTIFFS

Strategy: Use pattern matching to simplify names

- Convert to all caps
- Remove punctuation, common abbreviations (DBA, LLC, etc.)
- Standardize common words (APARTMENTS, MOBILE HOMES)
- Target frequently appearing names specifically

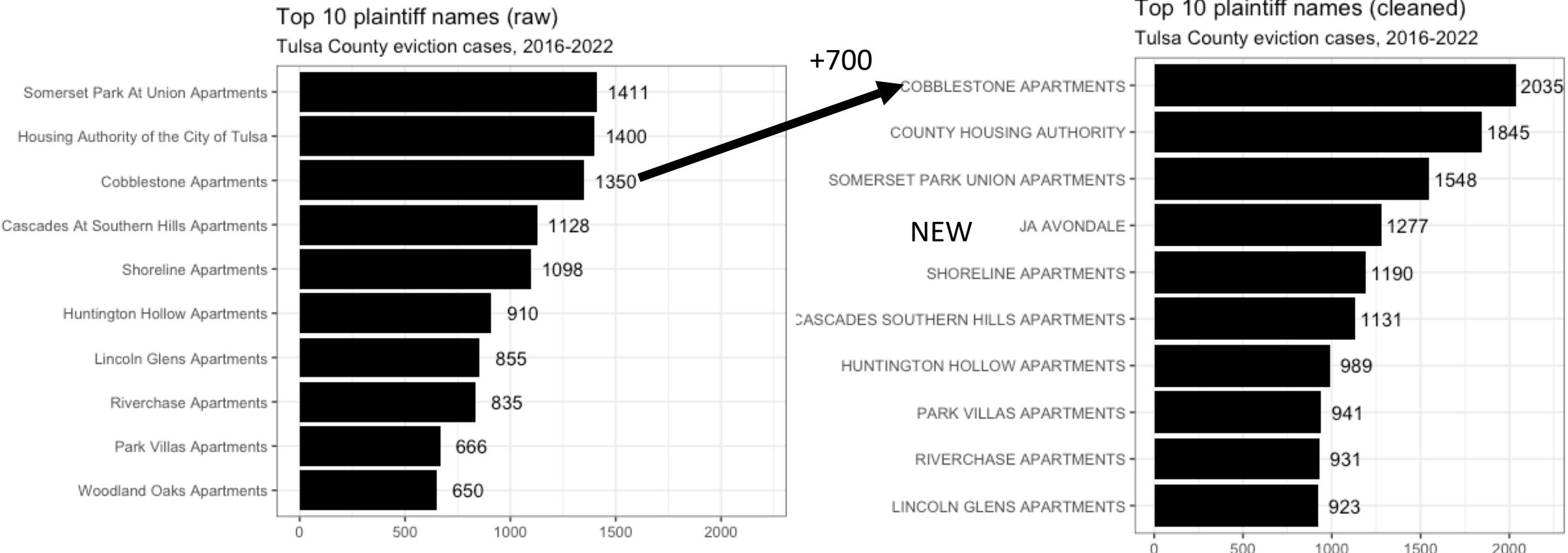


Title

```
1 clean_pl_names <- function(x) {  
2   str_to_upper({{x}}) |>  
3   str_remove_all( " INC| CO(?:=(\$| ))|LLC|LPII|DBA.*") |>  
4   str_replace_all("APT([[alpha:]])|APARTMENT([[alpha:]])$"  
5                   "APARTMENTS") |>  
6   str_replace("CHAT.*68.*", "CHATEAU 68 APARTMENTS")  
7 }
```

Code at bit.ly/ojo-at-gu

CLUSTERING AND CLASSIFYING: PLAINTIFFS



CLUSTERING AND CLASSIFYING: SERIAL EVICTION FILINGS

Question: Which plaintiffs are filing serial evictions against their tenants?

Problem: Defendant names are not standardized; filings by a landlord against same tenant may not always be “serial”

| | district | case_number | date_filed | plaint_clean | def |
|----|----------|---------------|------------|--------------------|-----|
| 1 | TULSA | SC-2017-465 | 2017-01-12 | MANSIONS RIVERSIDE | N C |
| 2 | TULSA | SC-2017-2576 | 2017-02-15 | MANSIONS RIVERSIDE | N C |
| 3 | TULSA | SC-2017-3933 | 2017-03-13 | MANSIONS RIVERSIDE | N C |
| 4 | TULSA | SC-2017-5466 | 2017-04-12 | MANSIONS RIVERSIDE | N C |
| 5 | TULSA | SC-2017-6987 | 2017-05-11 | MANSIONS RIVERSIDE | N C |
| 6 | TULSA | SC-2017-8764 | 2017-06-13 | MANSIONS RIVERSIDE | N C |
| 7 | TULSA | SC-2017-10323 | 2017-07-11 | MANSIONS RIVERSIDE | N C |
| 8 | TULSA | SC-2017-12378 | 2017-08-11 | MANSIONS RIVERSIDE | N C |
| 9 | TULSA | SC-2017-14270 | 2017-09-12 | MANSIONS RIVERSIDE | N C |
| 10 | TULSA | SC-2017-16153 | 2017-10-12 | MANSIONS RIVERSIDE | N C |
| 11 | TULSA | SC-2017-19737 | 2017-12-12 | MANSIONS RIVERSIDE | N C |
| 12 | TULSA | SC-2018-440 | 2018-01-11 | MANSIONS RIVERSIDE | N C |
| 13 | TULSA | SC-2018-2677 | 2018-02-14 | MANSIONS RIVERSIDE | N C |
| 14 | TULSA | SC-2018-4546 | 2018-03-14 | MANSIONS RIVERSIDE | N C |
| 15 | TULSA | SC-2018-5788 | 2018-04-11 | MANSIONS RIVERSIDE | N C |
| 16 | TULSA | SC-2018-9378 | 2018-06-12 | MANSIONS RIVERSIDE | N C |
| 17 | TULSA | SC-2018-11076 | 2018-07-12 | MANSIONS RIVERSIDE | N C |
| 18 | TULSA | SC-2018-13078 | 2018-08-13 | MANSIONS RIVERSIDE | N C |
| 19 | TULSA | SC-2018-14944 | 2018-09-11 | MANSIONS RIVERSIDE | N C |
| 20 | TULSA | SC-2018-16962 | 2018-10-11 | MANSIONS RIVERSIDE | N C |

CLUSTERING AND CLASSIFYING: SERIAL EVICTION FILINGS

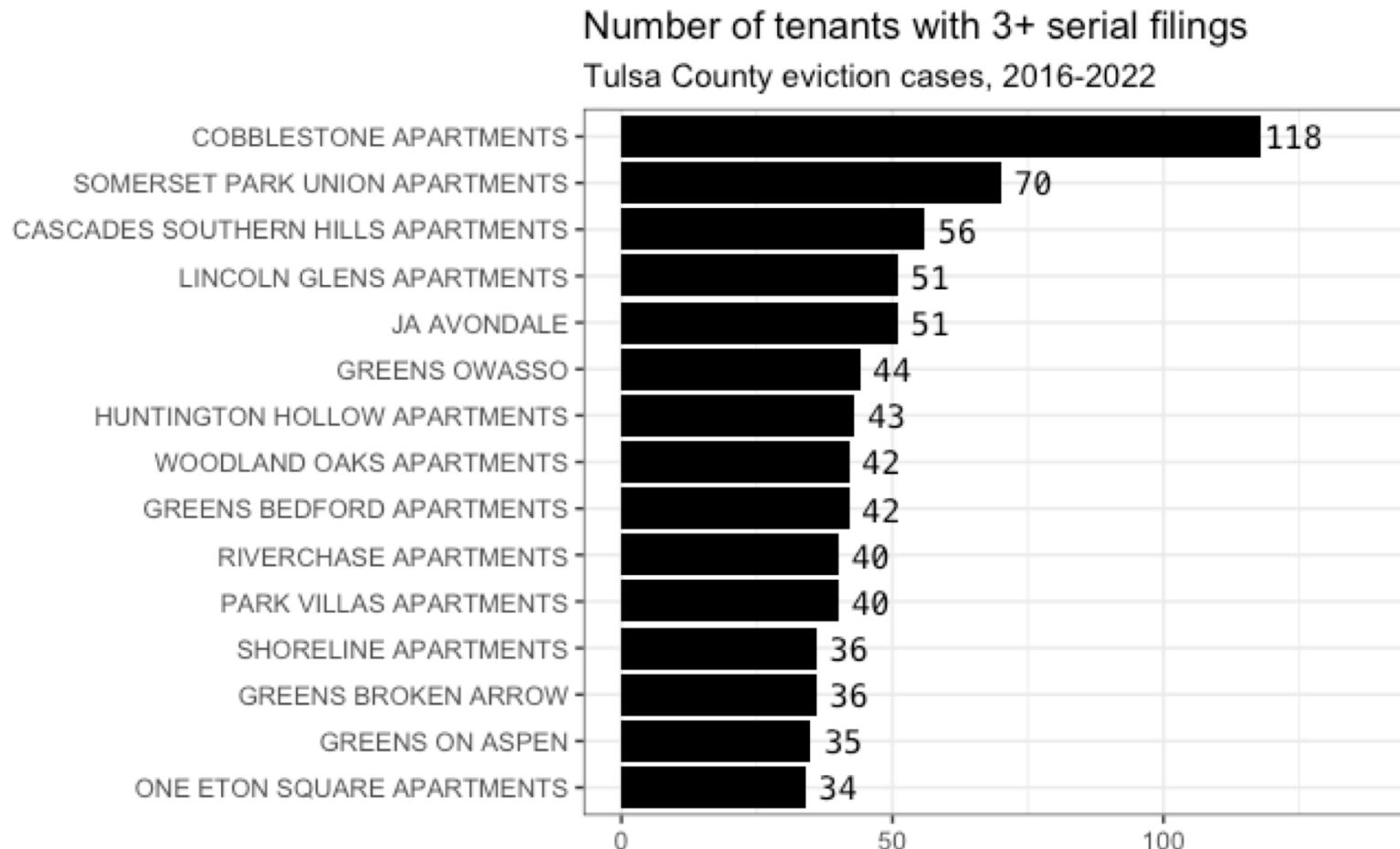
Strategy:

1. Eliminate middle initials from defendant names
2. Calculate days since last filing
3. Group by plaintiff and defendant
4. Choose cutoff for “serial” (we chose new filings within 70 days of previous)

```
● ● ● Title  
1 serial <- evictions |>  
2   group_by(plaint_clean, def) |>  
3   arrange(plaint_clean, def, date_filed) |>  
4   select(district, case_number, date_filed, plaint_clean, def) |>  
5   arrange(plaint_clean, def)  
6   mutate(days_since_last = as.numeric(date_filed) -  
       as.numeric(lag(date_filed))) |>  
7   filter(days_since_last < 70) |>  
8   count(plaint_clean, def)
```

Code at bit.ly/ojo-at-gu

CLUSTERING AND CLASSIFYING: SERIAL EVICTION FILINGS



CLUSTERING AND CLASSIFYING: EVICTION FILING TYPES

Question: Which eviction filings are for back rent and which are for lease violations?

Problem: Landlords say they often sue only for possession; service providers say that most are for past due rent. Court records do not indicate lease violations directly

| description.y |
|---|
| AMOUNT IN DEBT OF <..\$..>\$1,495.00 |
| AMOUNT IN DEBT OF <..\$..>\$1,774.00 |
| AMOUNT IN DEBT OF <..\$..>\$655.00 \$75.00 LATE FEE |
| AMOUNT IN DEBT OF \$1,137.00 |
| AMOUNT IN DEBT OF \$3,865.56 |
| AMOUNT IN DEBT OF \$975.00 |
| AMOUNT IN DEBT OF POSSESSION |
| AMOUNT IN DEBT OF - POSSESSION |
| AMOUNT IN DEBT OF <..\$..>\$545.00 \$75.00 LATE FEE |
| AMOUNT IN DEBT OF <..\$..>\$1,679.50 |

CLUSTERING AND CLASSIFYING: EVICTION FILING TYPES

Strategy:

1. Use regex to extract fees as list
2. Separate to columns
3. Convert to numeric and sum
4. Use regex to check if a late fee is noted

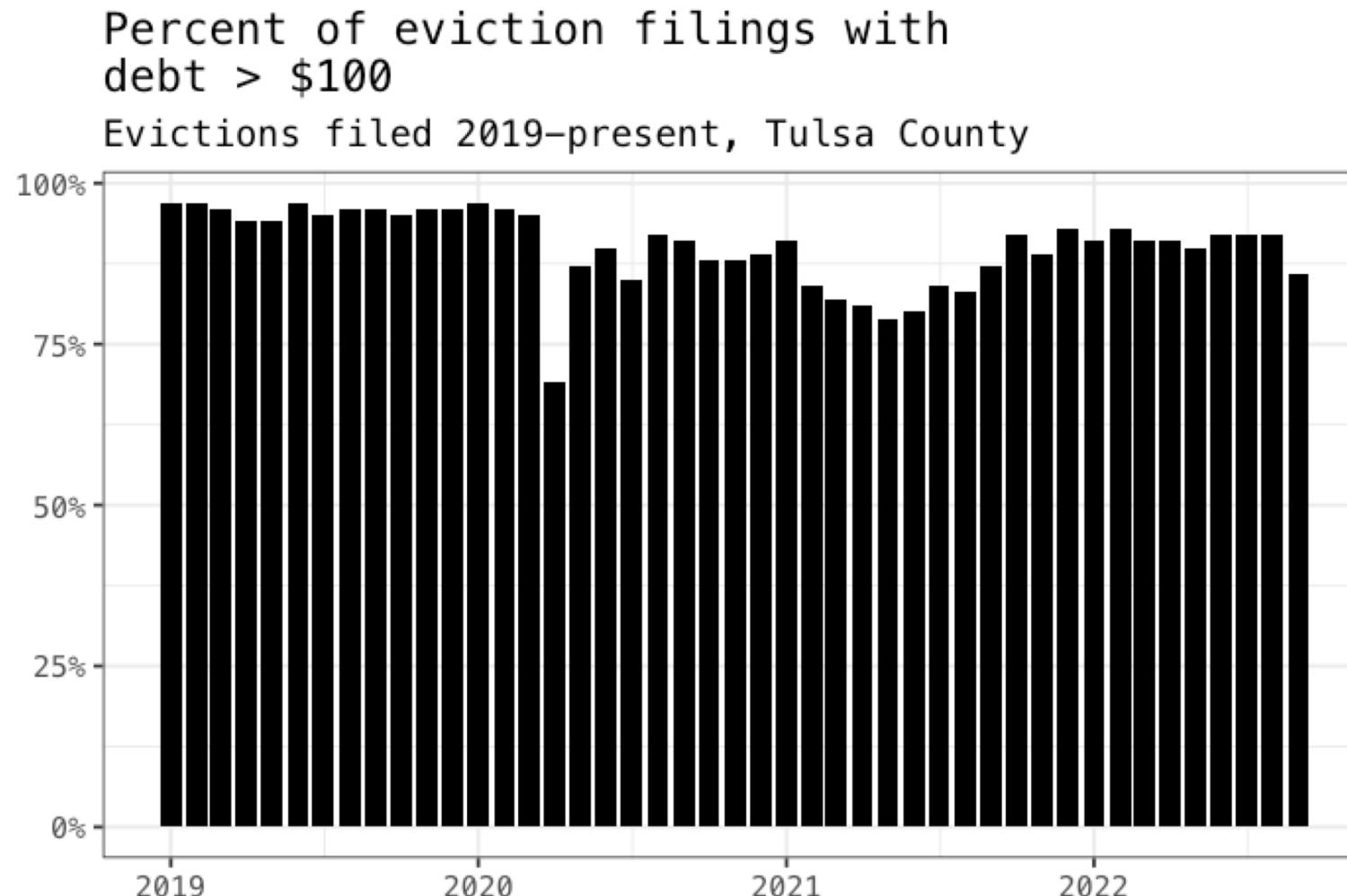
```
● ● ● Title

1 debt <- evictions %>%
2   select(district, case_number, min_desc = minute) %>%
3   distinct %>%
4   mutate(min = str_remove_all(min_desc, "AMOUNT IN DEBT OF
|\\$|,") %>%
5     str_remove_all("( PER |Document)(.|\\n)*") %>%
6     str_squish()) %>%
7   mutate(fees = str_extract_all(min, "(\\d|\\.\\.)+")) %>%
8     as.character() %>%
9     str_remove_all('c\\(\"|\"\')')) %>%
10  separate(fees,
11    into = paste0("fee", 1:4),
12    sep = ',', '') %>%
13  rowwise(district, case_number, min_desc) %>%
14  mutate(across(contains("fee"), as.numeric),
15    debt_amt = sum(feel, fee2, fee3, fee4, na.rm = TRUE),
16    late_fee = str_detect(min_desc, "LATE")) %>% # Add
      variable for late fee
17  group_by(district, case_number) %>%
18  summarize(debt_amt = sum(debt_amt),
19            late_fee = any(late_fee == TRUE))
```

Code at bit.ly/ojo-at-gu



CLUSTERING AND CLASSIFYING: EVICTION FILING TYPES



FINAL THOUGHTS

Advantages:

- Uses open source tools (no costly 3rd party services)
- Maintainable in local workflow (no uploading to a website, etc.)
- Can be refined over time
- Specific to your data

Disadvantages:

- Initial process is tedious
- Requires lots of patience and time
- Idiosyncrasies vary across jurisdictions



openjustice.okpolicy.org

RYAN GENTZLER
rgentzler@okpolicy.org