Leveraging vector processing with R to analyze StreetLight Data

An R package approach

METROPOLITAN COUNCIL

November 2022

Liz Roten

metrocouncil.org

Hello!

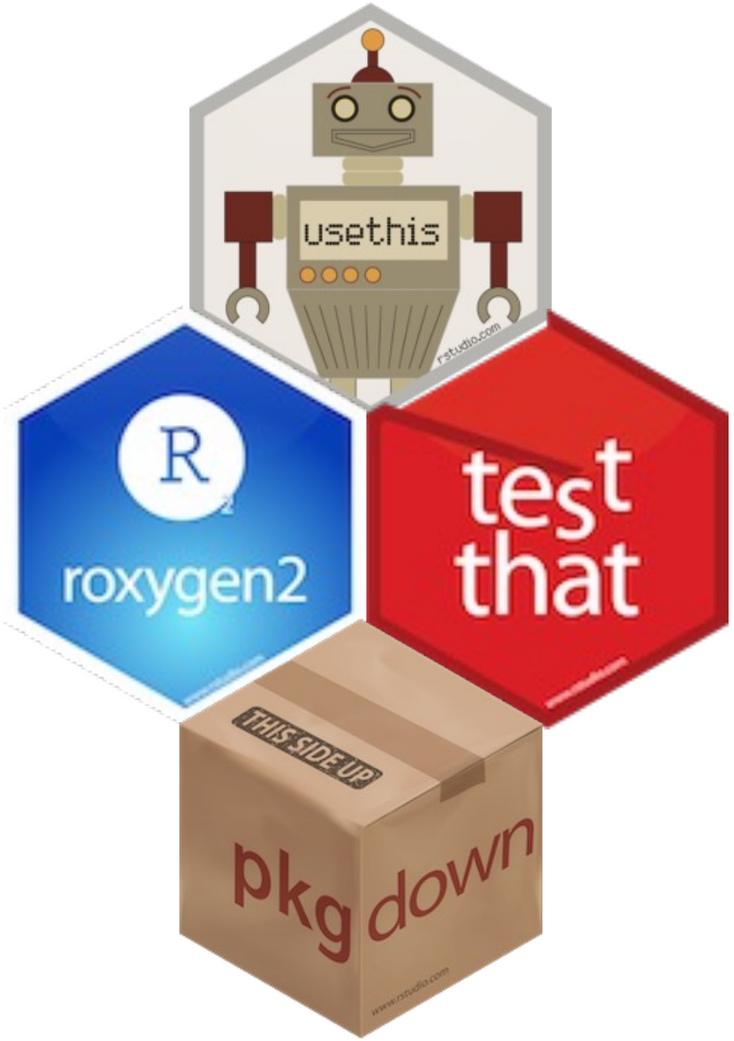
I’m Liz, and I’ve been using R professionally for ~ 4 years

- The #rstats community is small but mighty
- Culture of collaboration
- Genuine desire to help out

Why R?

And why an R package?

- Plentiful resources
- Version control
- Collaboration
- Import StreetLight data directly into analysis



Enter {streetlightR}

R wrapper for StreetLight API

- Available on GitHub!
- Nice error and warning messages that make coding more delightful
- Automated testing on Mac and Windows
- Documentation
- Long and short-form examples

Using {streetlightR}

Great opportunity for vector-based programming

- Less time spent on memory allocation
- No side effects – user environment is preserved
- Write less and more efficient code
- Option for parallel computing

Case Study
Pandemic Vehicle Miles Traveled and Trip Length

Research questions

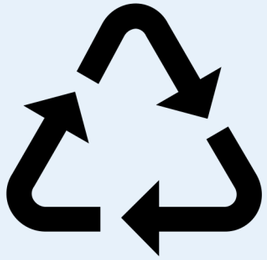
Did the pandemic have a quantifiable impact on VMT?

- How did trip length change over the last three years?
- Are there seasonal trends?
- Are there differences in urban vs. rural areas?

Parameters

Static, reusable

- Analysis type
- Day types
- Day parts
- Zones
- Output type
- Travel mode
- Add-on’s



Dynamic

- Start and end date
- Analysis name

Create month sequence

start_date	end_date	month_year	month_of_year
01/01/2019	01/31/2019	2019	1
02/01/2019	02/28/2019	2019	2
03/01/2019	03/31/2019	2019	3
04/01/2019	04/30/2019	2019	4
05/01/2019	05/31/2019	2019	5
06/01/2019	06/30/2019	2019	6
07/01/2019	07/31/2019	2019	7
...	...	...	...

Code and run

```
tictoc::tic(msg = "Create volume analyses")
volume_monthly <- purrr::map(
  .x = c(1:nrow(month_sequence)),
  function(analysis_month){
    Sys.sleep(10)

    create_streetlight_analysis(
      login_email = login_email,
      api_key = api_key,
      analysis_type = "Zone_Activity_Analysis",
      analysis_name = paste0( "v", vers,
                              "-VMT ",
                              month_sequence[analysis_month,]$month_year, " month
",                            month_sequence[analysis_month,]$month_of_year,
                              " Volume Trip Trav Attr"),
      travel_mode_type = "All_Vehicles",
      output_type = "volume",
      origin_zone_set = "MPO_Counties",
      trip_attributes = TRUE,
      traveler_attributes = TRUE,
      date_ranges = list(start_date = month_sequence[analysis_month,]$start_date,
                         end_date = month_sequence[analysis_month,]$end_date),
      day_types = day_types,
      day_parts = day_parts,
      tags = list("streetlightR")
    )
  })
tictoc::toc()
```

1 zone set

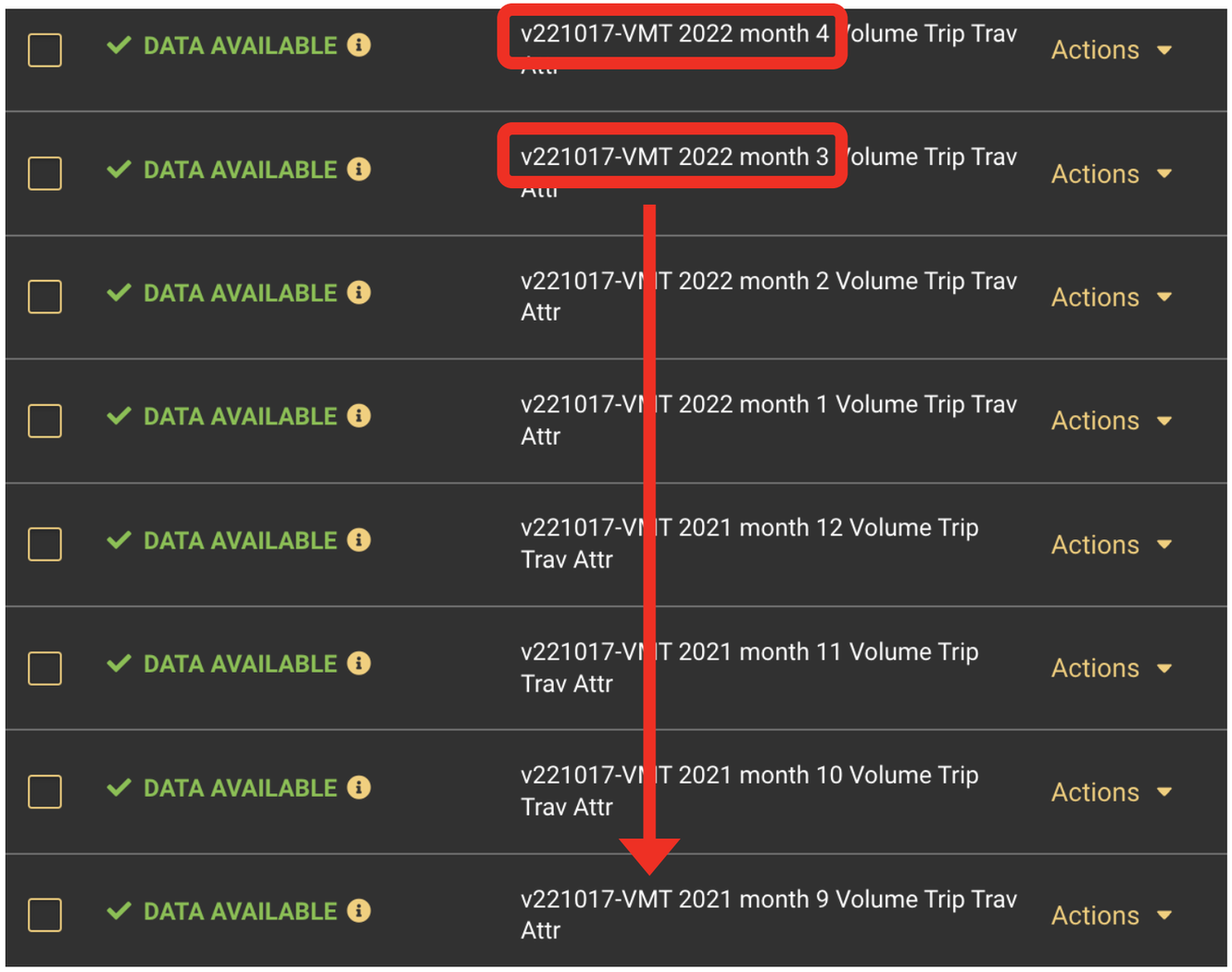
40 time periods

= 40 analyses

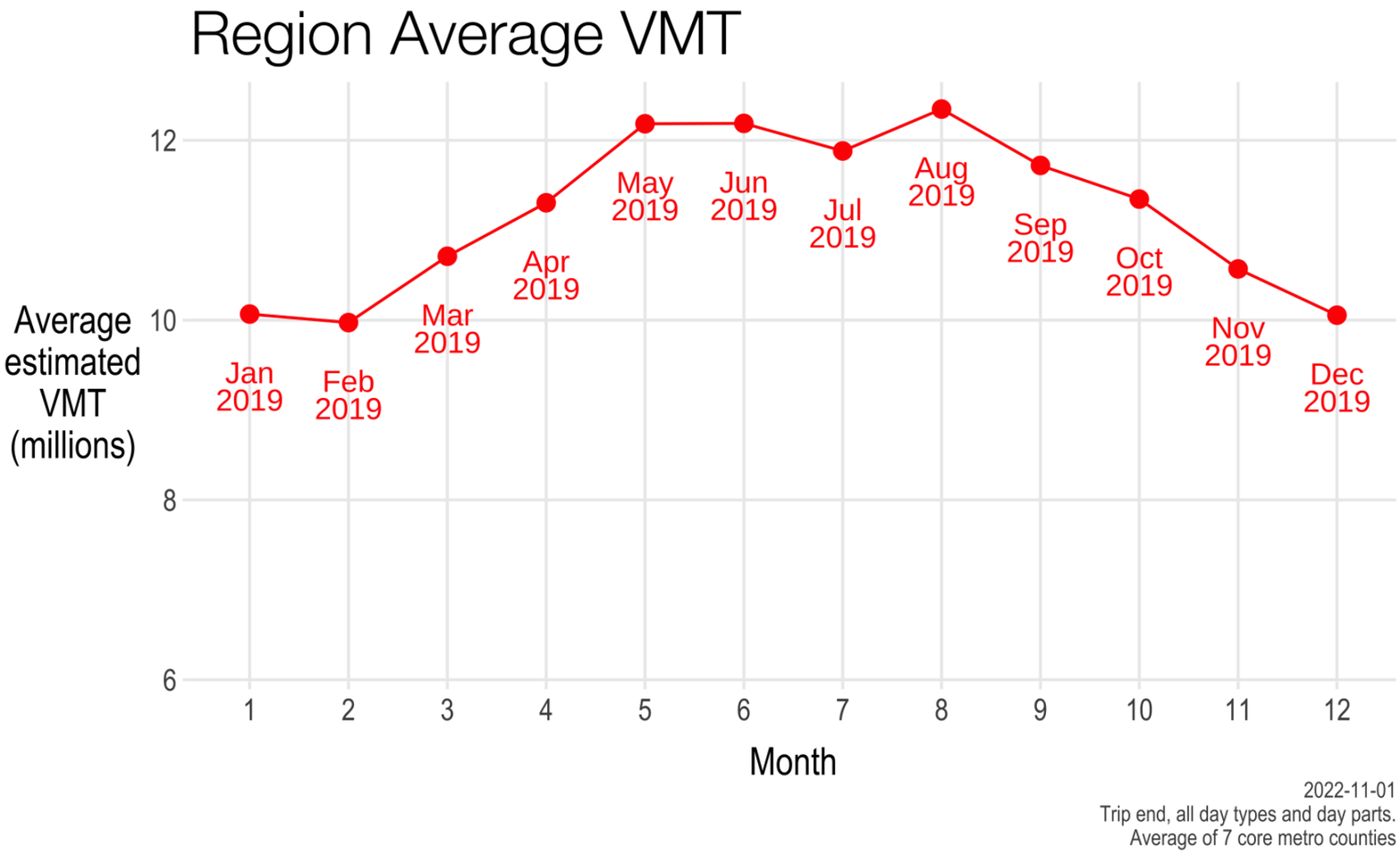
Re-usable: just complete a fresh run once a month

~ 3-hour runtime

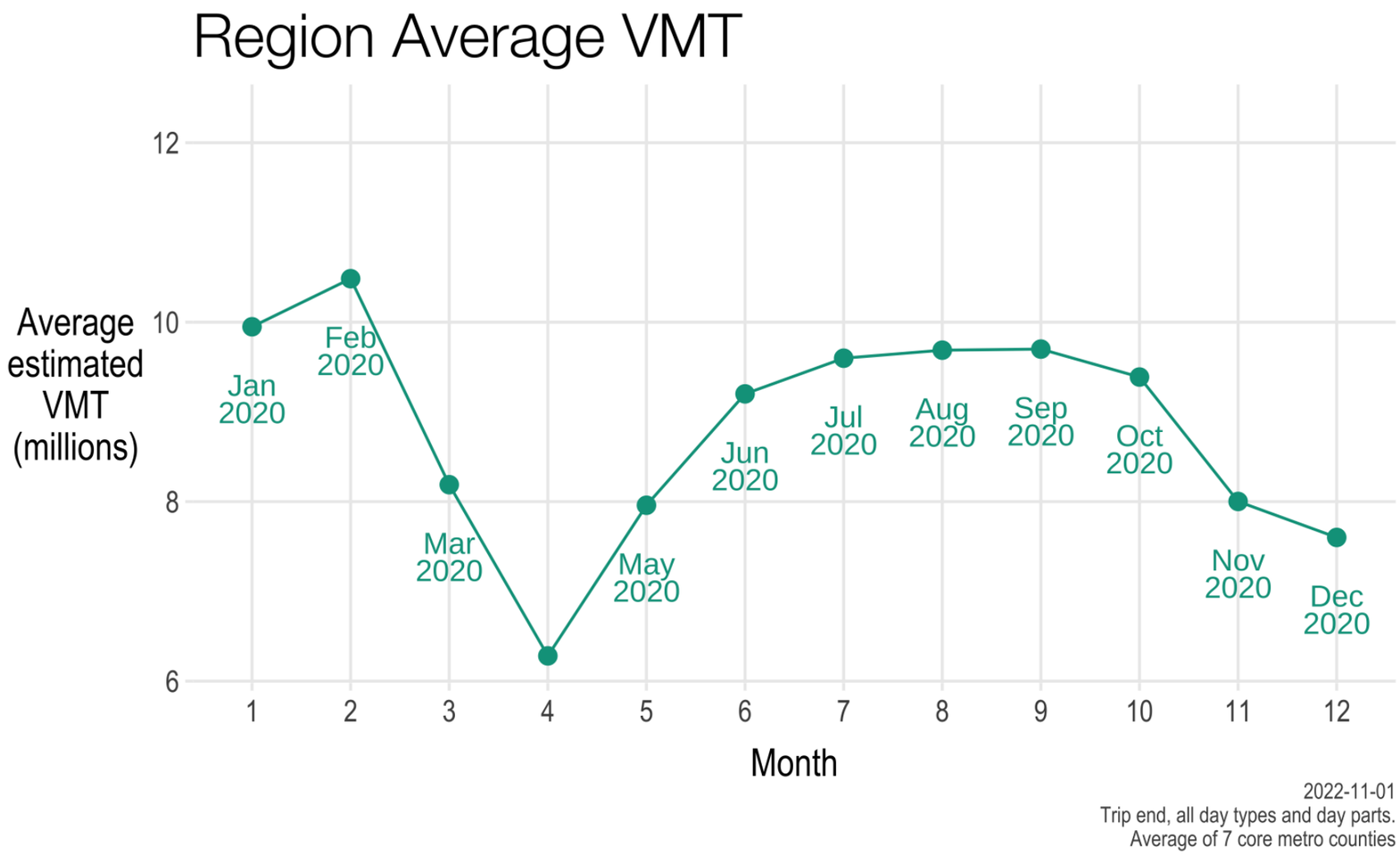
In InSights®


DATA AVAILABLE
v221017-VMT 2022 month 4 Volume Trip Trav Attr
Actions
DATA AVAILABLE
v221017-VMT 2022 month 3 Volume Trip Trav Attr
Actions
DATA AVAILABLE
v221017-VMT 2022 month 2 Volume Trip Trav Attr
Actions
DATA AVAILABLE
v221017-VMT 2022 month 1 Volume Trip Trav Attr
Actions
DATA AVAILABLE
v221017-VMT 2021 month 12 Volume Trip Trav Attr
Actions
DATA AVAILABLE
v221017-VMT 2021 month 11 Volume Trip Trav Attr
Actions
DATA AVAILABLE
v221017-VMT 2021 month 10 Volume Trip Trav Attr
Actions
DATA AVAILABLE
v221017-VMT 2021 month 9 Volume Trip Trav Attr
Actions

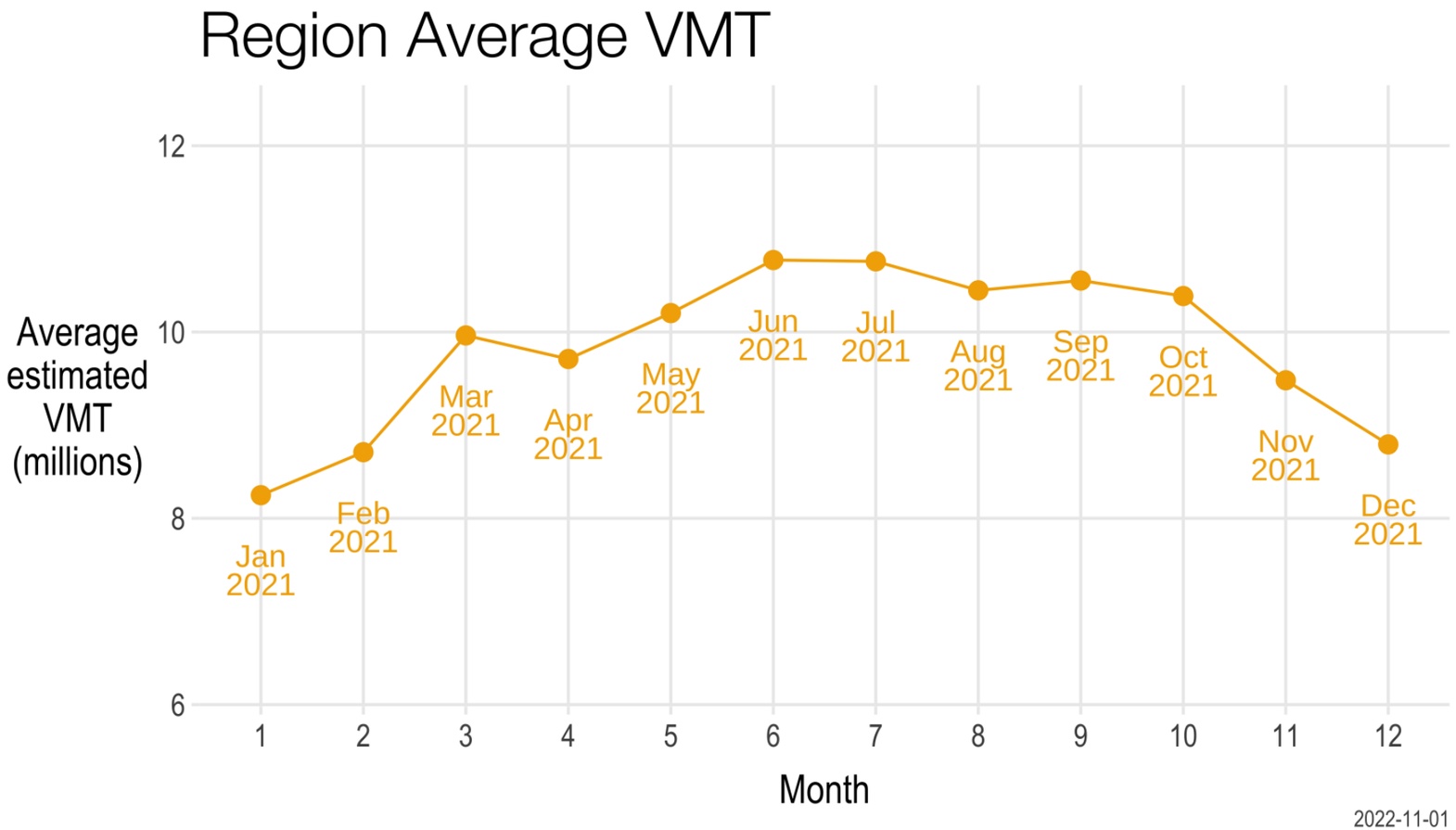
VMT 2019


Region Average VMT
Average estimated VMT (millions)
12
10
8
6
1
2
3
4
5
6
7
8
9
10
11
12
Month
Jan 2019
Feb 2019
Mar 2019
Apr 2019
May 2019
Jun 2019
Jul 2019
Aug 2019
Sep 2019
Oct 2019
Nov 2019
Dec 2019
2022-11-01
Trip end, all day types and day parts.
Average of 7 core metro counties

VMT 2020

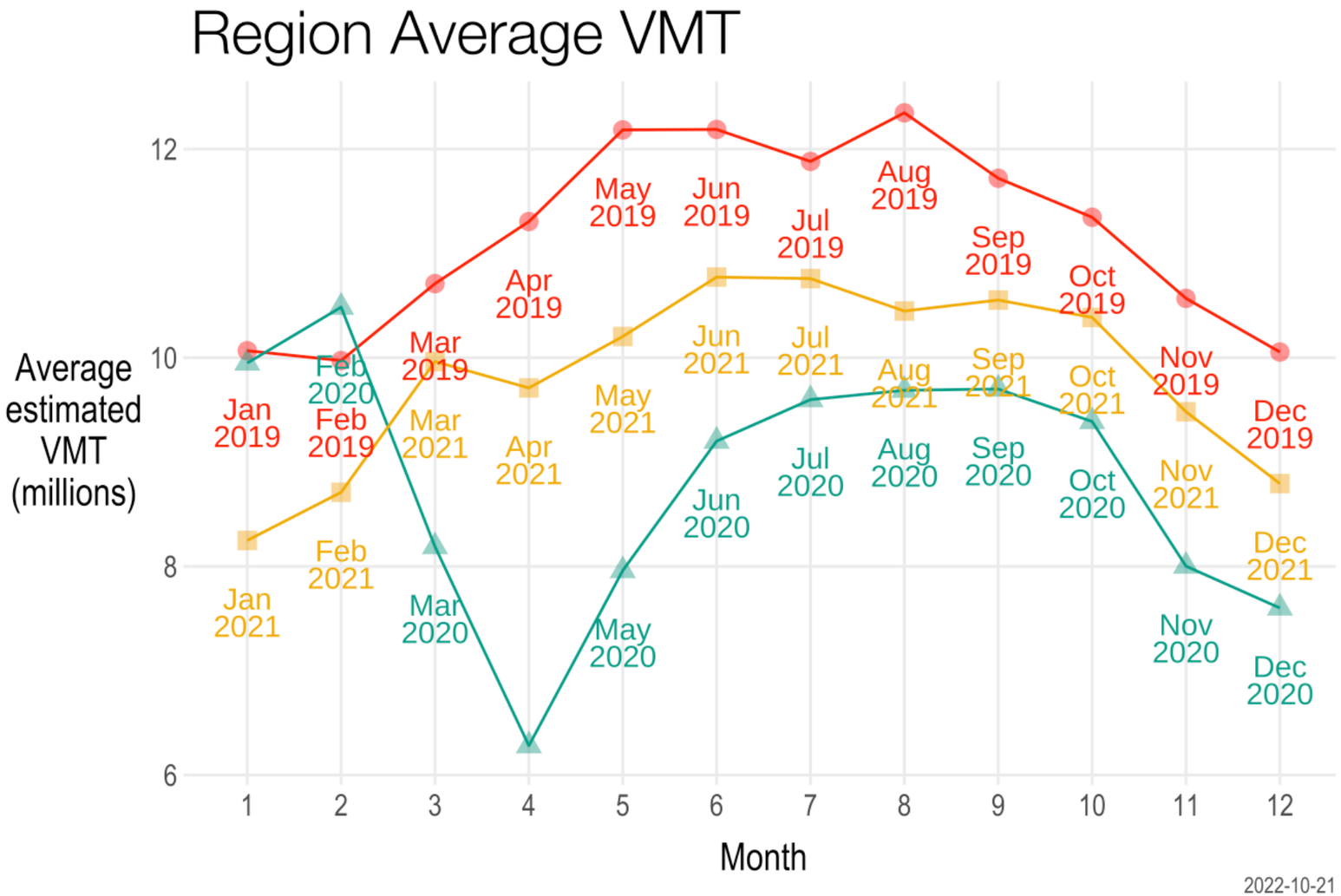

Region Average VMT
Average estimated VMT (millions)
12
10
8
6
1
2
3
4
5
6
7
8
9
10
11
12
Month
Jan 2020
Feb 2020
Mar 2020
May 2020
Jun 2020
Jul 2020
Aug 2020
Sep 2020
Oct 2020
Nov 2020
Dec 2020
2022-11-01
Trip end, all day types and day parts.
Average of 7 core metro counties

VMT 2021



2022-11-01
Trip end, all day types and day parts.
Average of 7 core metro counties

VMT 2019-21



2022-10-21
Trip end, all day types and day parts.
Average of 7 core metro counties

Case Study
Regional and state parks research

Context

Joint project with 3 agencies to research parks and LBS data

DEPARTMENT OF NATURAL RESOURCES

GREATER MINNESOTA REGIONAL Parks and Trails COMMISSION

METROPOLITAN COUNCIL

Park Research questions

- Can location-based services measure park and trail visitation across the state of Minnesota?
- How does visitation differ between parks?
    - How did COVID affect park visitation?
- What can we learn about visitor demographics?
- What is the mode share for each park?

Scale

1 zone set (225+ zones)

52 weeks/year

3 years

2 analysis types

3 modes

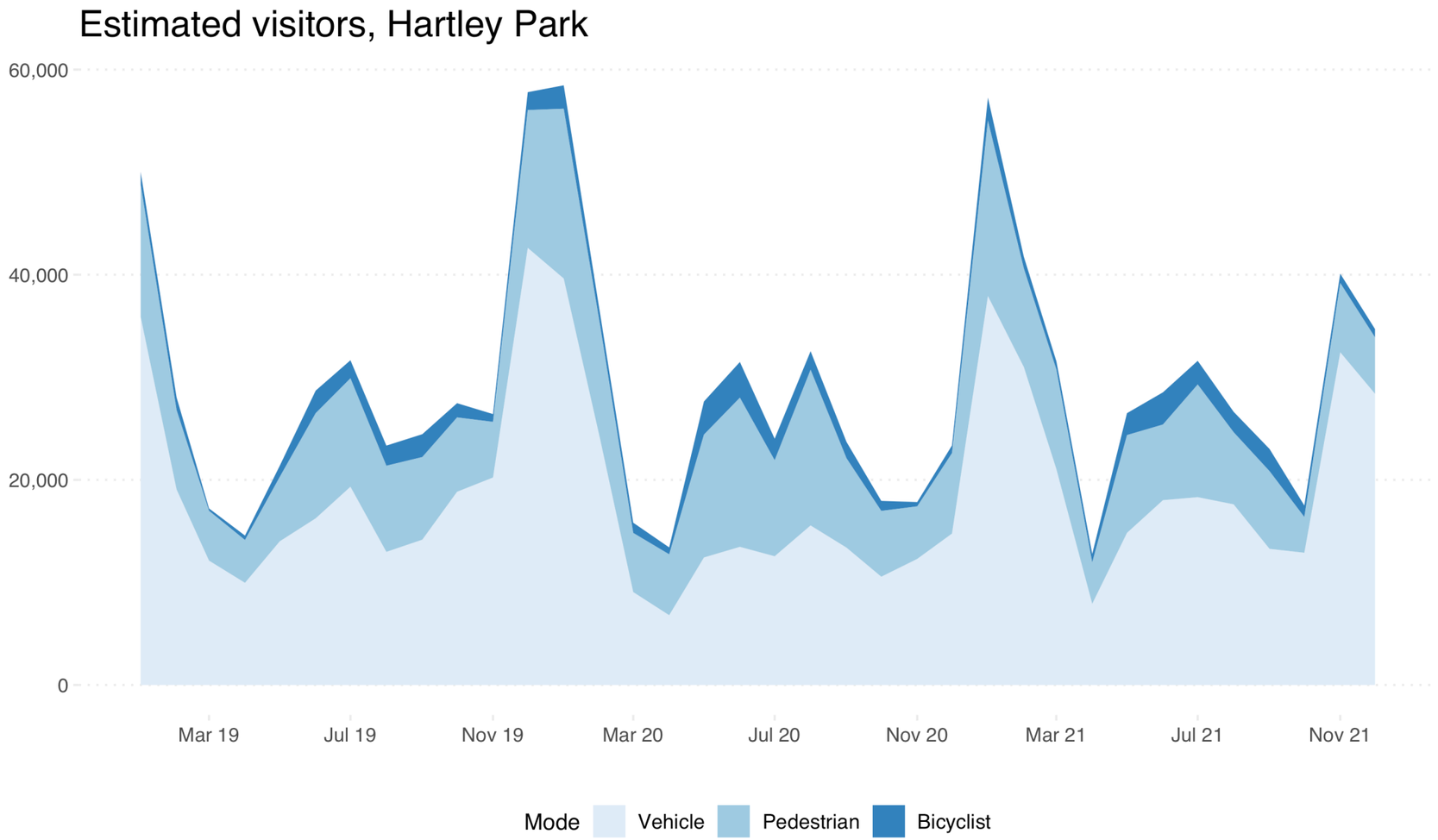
= 936 analyses

Plus, iterations, releases, and experimentation

Between June 1, 2021, and November 1, 2022

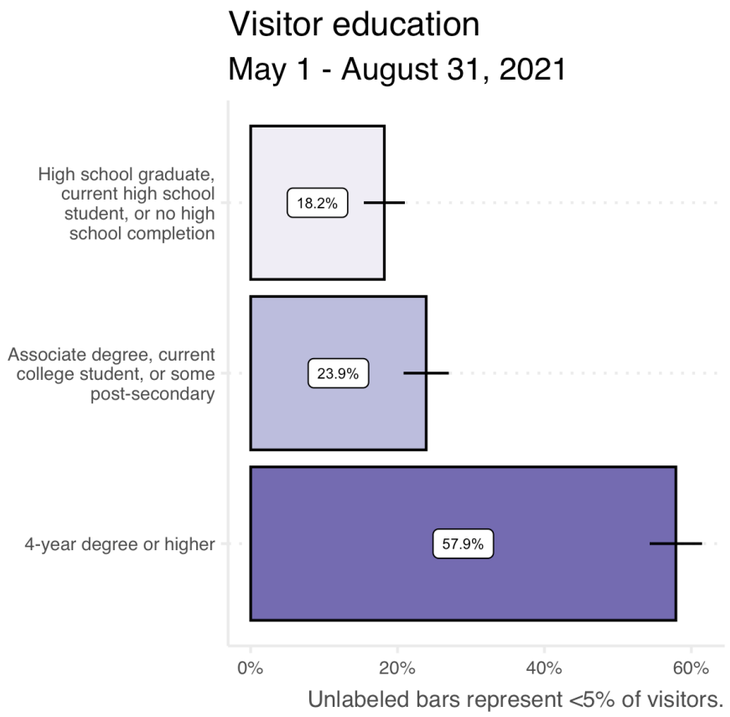
Team member	Analyses
RM	5,217
EE	1,946

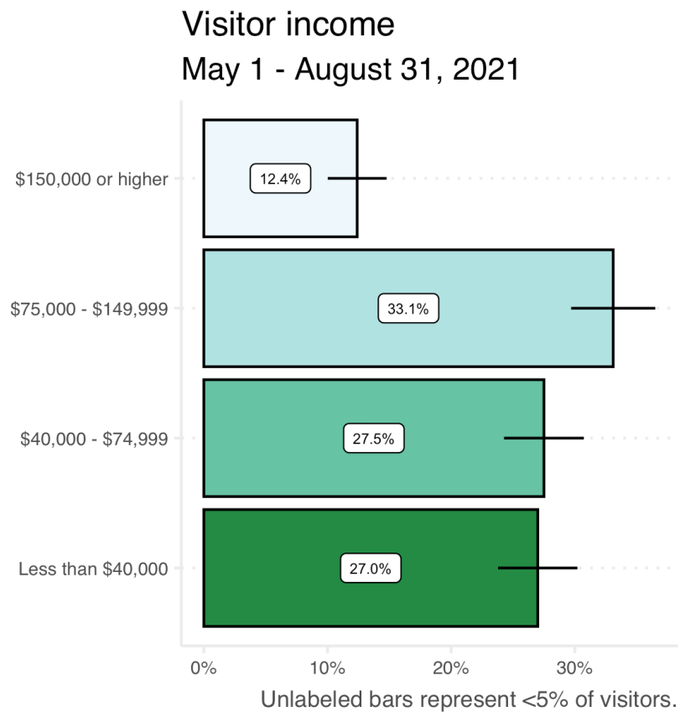
Results – Unprecedented detail

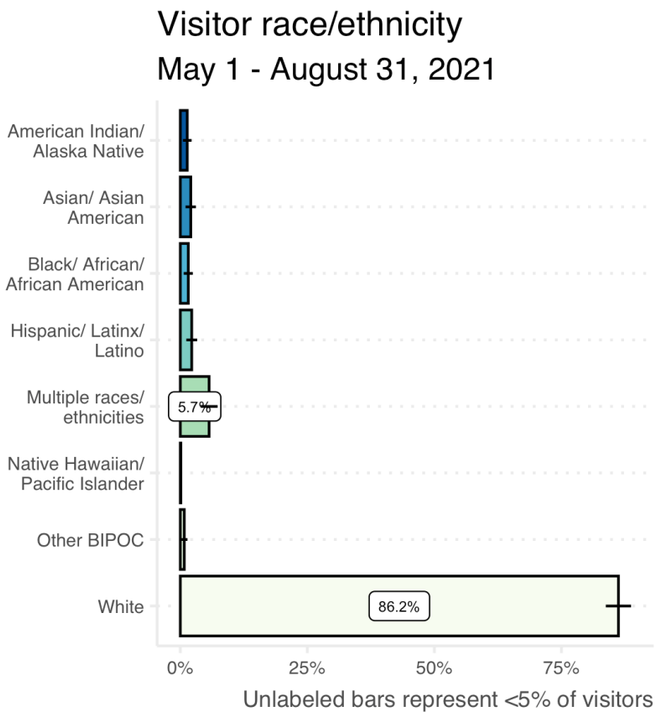


Source: StreetLight Data, accessed August 2022.

Results – Visitor demographics







Source: StreetLight Data. Accessed August 2022.

This Project was funded with Legacy Partnership Research Funds from the State of Minnesota Parks and Trails Legacy Fund

More case studies

Questions we didn't have before {streetlightR}

- One agency removed parking fees – did that affect visitation?
- Another recently reintroduced buffalo – what is the before/after affect?
- When should we send out surveyors to a boat launch if we have limited resources?

Liz Roten

Data Scientist, CD/MTS
liz.roten@metc.state.mn.us

Ellen Esch, Ph.D.

Senior Data Scientist, CD
ellen.esch@metc.state.mn.us

Raven McKnight

Associate Data Scientist, CD
raven.mcknight@metc.state.mn.us

METROPOLITAN
COUNCIL