Tut_2_Code

AUTHOR
Ricky Fung

Preamble

Purpose: To examine the number of reported crimes based on the category of crime in the year

2014 and create a graph.

Date: 2024/01/15

Contact: ricky.fung@mail.utoronto.ca

Workspace setup

load packages

```
library(janitor)
library(tidyverse)
library(dplyr)
library(opendatatoronto)
```

get package

```
package <- show_package("police-annual-statistical-report-reported-crimes")
```

get all resources for this package

```
resources <- list_package_resources("police-annual-statistical-report-reported-crimes")
```

identify datastore resources; by default, Toronto Open Data sets datastore resource format to CSV for non-geospatial and GeoJSON for geospatial resources

```
datastore_resources <- filter(resources, tolower(format) %in% c('csv', 'geojson'))
```

load the first datastore resource as a sample

```
data <- filter(datastore_resources, row_number()==1) %>% get_resource()
head(data)
```

```
# A tibble: 6 × 7
  `_id` REPORT_YEAR DIVISION CATEGORY               SUBTYPE COUNT_ COUNT_CLEARED
  <int>       <int> <chr>    <chr>                  <chr>   <chr>  <chr>
1     1        2014 D11      Crimes Against the Pe… Other   22     9
2     2        2014 D11      Crimes Against Proper… Theft … 1      1
3     3        2014 D11      Crimes Against the Pe… Other   1      1
4     4        2014 D11      Crimes Against the Pe… Robber… 1      1
5     5        2014 D11      Crimes Against Proper… Break … 23     13
6     6        2014 D11      Crimes Against Proper… Theft … 1      1
```

Cleaning and tidying data

Clean column names

```
cleaned <-
  clean_names(data)
```

Select necessary columns

```
cleaned <-
  cleaned |>
  select(
    id,
    category
  )

head(cleaned)
```

```
# A tibble: 6 × 2
     id category
  <int> <chr>
1     1 Crimes Against the Person
2     2 Crimes Against Property
3     3 Crimes Against the Person
4     4 Crimes Against the Person
```

```
5     5 Crimes Against Property
6     6 Crimes Against Property
```

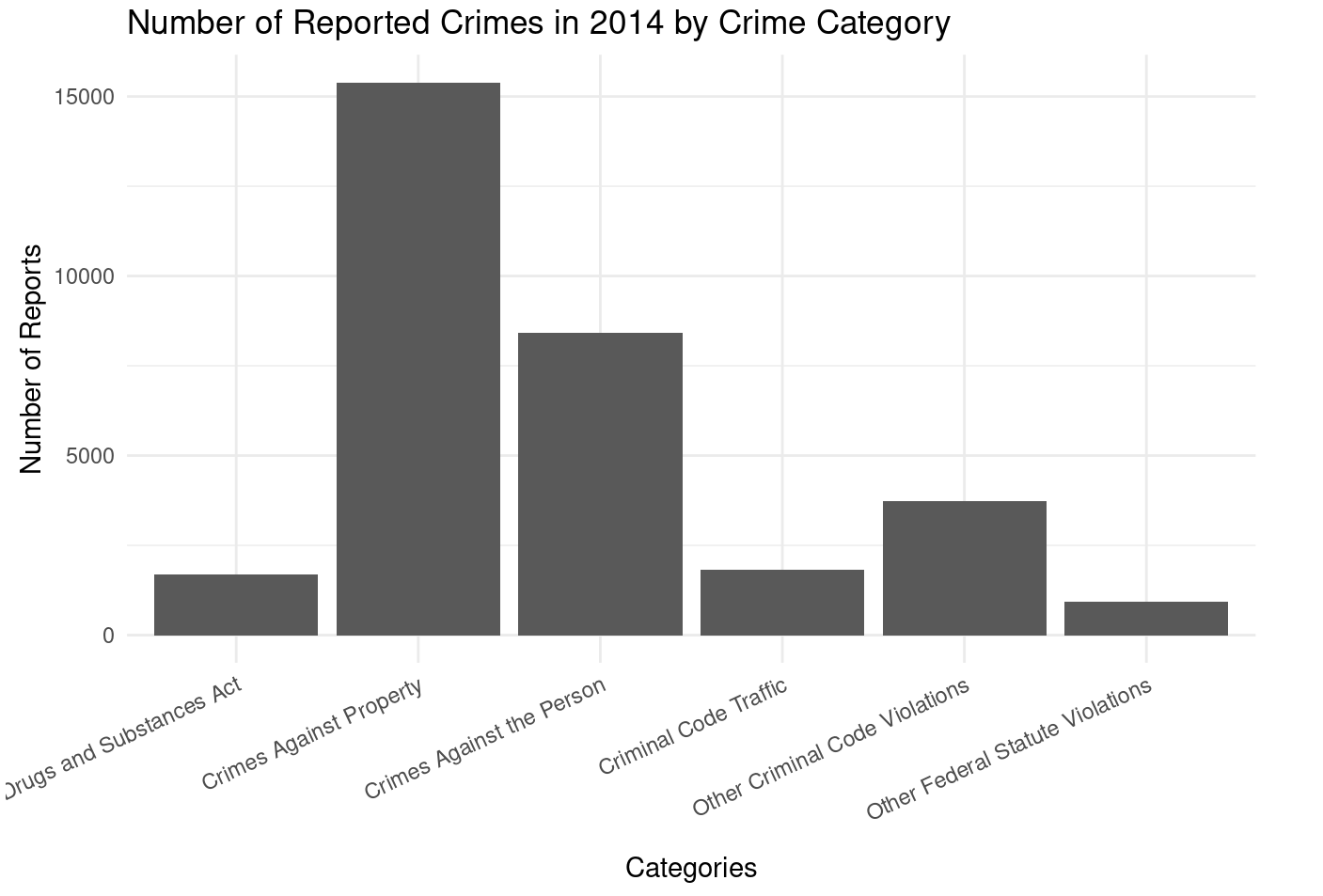
Save cleaned data as csv file

```
write_csv(x=cleaned, file="reported_crimes_reports.csv")
```

Plot

Create bar graph to find the number of reported crimes for each category in 2014

```
cleaned |>
  ggplot(aes(x = category)) +
  geom_bar() +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 25, hjust = 1)) +  # Rotate x-axis labels
  labs(title = "Number of Reported Crimes in 2014 by Crime Category", x = "Categories", y = "Numb
```



```
cleaned |>
  count(category)
```

```
# A tibble: 6 × 2
  category                              n
  <chr>                             <int>
1 Controlled Drugs and Substances Act  1696
2 Crimes Against Property             15392
3 Crimes Against the Person            8423
4 Criminal Code Traffic                1832
5 Other Criminal Code Violations       3737
6 Other Federal Statute Violations      920
```

Share

The 6 categories of crime we are interested in are Controlled Drugs and Substances Act, Crimes Against Property, Crimes Against the Person, Criminal Code Traffic, Other Criminal Code Violations, Other Federal Statute Violations. We are interested in the number of reports made for each category.

We found that the crimes: Controlled Drugs and Substances Act had 1696 reports, Crimes Against Property had 15392, Crimes Against the Person had 8423, Criminal Code Traffic had 1832, Other Criminal Code Violations had 3737, and Other Federal Statute Violations had 920 reports.

Citations

```
citation()
```

```
To cite R in publications use:

  R Core Team (2023). _R: A Language and Environment for Statistical
  Computing_. R Foundation for Statistical Computing, Vienna, Austria.
  <https://www.R-project.org/>.

A BibTeX entry for LaTeX users is

  @Manual{,
    title = {R: A Language and Environment for Statistical Computing},
    author = {{R Core Team}},
    organization = {R Foundation for Statistical Computing},
    address = {Vienna, Austria},
    year = {2023},
    url = {https://www.R-project.org/},
  }

We have invested a lot of time and effort in creating R, please cite it
```

```
when using it for data analysis. See also 'citation("pkgname")' for
citing R packages.
```

```
citation("janitor")
```

```
To cite package 'janitor' in publications use:

  Firke S (2023). _janitor: Simple Tools for Examining and Cleaning
  Dirty Data_. R package version 2.2.0,
  https://sfirke.github.io/janitor/,
  <https://github.com/sfirke/janitor>.

A BibTeX entry for LaTeX users is

  @Manual{,
    title = {janitor: Simple Tools for Examining and Cleaning Dirty Data},
    author = {Sam Firke},
    year = {2023},
    note = {R package version 2.2.0,
https://sfirke.github.io/janitor/},
    url = {https://github.com/sfirke/janitor},
  }
```

```
citation("tidyverse")
```

```
To cite package 'tidyverse' in publications use:

  Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R,
  Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller
  E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V,
  Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H (2019). "Welcome to
  the tidyverse." _Journal of Open Source Software_, *4*(43), 1686.
  doi:10.21105/joss.01686 <https://doi.org/10.21105/joss.01686>.

A BibTeX entry for LaTeX users is

  @Article{,
    title = {Welcome to the {tidyverse}},
    author = {Hadley Wickham and Mara Averick and Jennifer Bryan and Winston Chang and Lucy
D'Agostino McGowan and Romain François and Garrett Grolemund and Alex Hayes and Lionel Henry and
Jim Hester and Max Kuhn and Thomas Lin Pedersen and Evan Miller and Stephan Milton Bache and
Kirill Müller and Jeroen Ooms and David Robinson and Dana Paige Seidel and Vitalie Spinu and
Kohske Takahashi and Davis Vaughan and Claus Wilke and Kara Woo and Hiroaki Yutani},
    year = {2019},
    journal = {Journal of Open Source Software},
    volume = {4},
    number = {43},
    pages = {1686},
```

```
    doi = {10.21105/joss.01686},
  }
```

```
citation("dplyr")
```

```
To cite package 'dplyr' in publications use:

  Wickham H, François R, Henry L, Müller K, Vaughan D (2023). _dplyr: A
  Grammar of Data Manipulation_. R package version 1.1.4,
  https://github.com/tidyverse/dplyr, <https://dplyr.tidyverse.org>.

A BibTeX entry for LaTeX users is

  @Manual{,
    title = {dplyr: A Grammar of Data Manipulation},
    author = {Hadley Wickham and Romain François and Lionel Henry and Kirill Müller and Davis
Vaughan},
    year = {2023},
    note = {R package version 1.1.4, https://github.com/tidyverse/dplyr},
    url = {https://dplyr.tidyverse.org},
  }
```

```
citation("opendatatoronto")
```

```
To cite package 'opendatatoronto' in publications use:

  Gelfand S (2022). _opendatatoronto: Access the City of Toronto Open
  Data Portal_. R package version 0.1.5,
  https://github.com/sharlagelfand/opendatatoronto/,
  <https://sharlagelfand.github.io/opendatatoronto/>.

A BibTeX entry for LaTeX users is

  @Manual{,
    title = {opendatatoronto: Access the City of Toronto Open Data Portal},
    author = {Sharla Gelfand},
    year = {2022},
    note = {R package version 0.1.5,
https://github.com/sharlagelfand/opendatatoronto/},
    url = {https://sharlagelfand.github.io/opendatatoronto/},
  }
```