Toronto's Bike Share program is one size but does not fit all*

Annual and casual riders are have differentiated needs and use cases but are served the same Bike Share experience

Hudson Yuen

06 February 2022

Abstract

Bike Share programs bring clear health and environmental benefits to cities like Toronto, making them an important part of a sustainable urban future. This report examines ridership data to explore the current state of Toronto's Bike Share program, aiming to identify how the experience can be improved for current and prospective riders alike. Existing behaviour is strongly differentiated between annual and casual riders, opening up opportunities to better serve each segment and grow the rider base in the process. The casual segment in particular can benefit from increased ride duration limits and potential location-specific investments.

1 Introduction

Cities are an integral part of a sustainable future, with transportation being a key area of improvement towards reaching net zero carbon emissions. Urban mobility is a critical aspect of this transition; road traffic accounts for around 80% of carbon emissions from transportation (Dia et al. 2019). Bikeshare programs have therefore been touted as one of many tools that cities to employ, taking cars off the road while providing other mental and physical health benefits (Schwanen 2021). Toronto has run a substantial bikeshare network since 2011 (hereafter referred to as "Bike Share"), but also lowly amongst Canadian cities in terms of the percentage of total population living within bike sharing service areas (Hosford and Winters 2018).

The data surrounding Bike Share is indispensable in helping improve the program and other pieces of sustainable urban infrastructure. Understanding who uses the system, for what purposes, and in what ways can help the city improve the experience and attract more riders. Both annual and casual riders are integral to the vitality of the network; however, they have substantially different use cases that are arguably not well-served by the current "one-size-fits-all" approach. There is little differentiation between the 2 ridership segments other than pricing, leading to casual riders facing unnecessary pain points and potentially discouraging further use.

There is immense growth potential with Bike Share, given that only 18% of Toronto's population were within service areas as recently as 2018 (Hosford and Winters 2018). Improving the current experience could greatly aid update of the program, whether that be through new initiatives - such as lockers for commuters - or more lenient time limitations for day-trippers. It is imperative that such improvements are made with consideration of two distinct rider segments, as serving both groups is key to building a strong and sustainable program.

The remainder of this paper is structured as follows: Section 2 covers the data source and analysis methodology employed. Section 3 then examines key dimensions of the data, proposing conclusions and areas of further exploration.

*All data and scripts available at https://github.com/hudyu17/bikeshare

2 Data

2.1 Data Source

This report examines Toronto Bike Share trip data from June to December 2021. This anonymised trip data is provided by the Toronto Parking Authority (TPA), the organisation behind Toronto's iconic Green P parking infrastructure. TPA has provided trip data to the public through the City of Toronto Open Data Portal since 2014 and on an ongoing basis. This dataset was last updated in December 2021, and the `R` package `opendatatoronto` (Gelfand 2020) was used to obtain all relevant data.

2.2 Methodology and Data Collection

Barring errors, the original dataset covers all rides undertaken on the Toronto Bike Share network since 2014. A welcome aspect for analysis is the mechanical nature of this data; trips are automatically added to the dataset when rides begin and end, obviating the need for self-reporting or other unreliable sources. In other words, there are theoretically no non-respondents. Each ride is associated with the following features: trip duration, start and end time, start and end station, bike ID, and user type (casual or annual member).

Note that extracting data across multiple years is cumbersome, and the months of January to May 2021 feature large swaths of missing data. It was also difficult to match station IDs to geographical coordinates, requiring joins with data from other APIs that proved unreliable.

`R` (R Core Team 2021) was the language and environment used for the bulk of this analysis, alongside the `tidyverse` (Wickham et al. 2019), `janitor` (Firke et al. 2021), `dplyr` (Wickham et al. 2021), `kableExtra` (Zhu et al. 2021), and `data table` (Dowle et al. 2021) packages. My workflow began with the `data-cleaning.R` script; data was downloaded using the `opendatatoronto` package and cleaned as follows:

- Renaming columns - e.g. "Trip..Duration" to "trip_duration"
- Removing outliers - trips less than 2 minutes or more than 37 hours in length, the latter based on Q3 + 1.5 x IQR
- Populating `NULL` station names based on existing name/ID pairs in the dataset
- Removing other `NULL` values

The result was saved to `toronto-bikes-clean.csv` and brought into `data-exploration.Rmd`. Exploratory analysis revealed interesting summary differences when segmenting by user types, forming the basis for further investigation.

2.3 Data Characteristics

The initial cleaned dataset shows data associated with each ride; a subset of rows and columns can be seen in Table 1 below. This analysis dataset contains anonymised trip data from approximately 2.53 million rides across 7 months of activity.

Table 1: Subset of initial dataset

trip_duration	start_station_name	user_type
831	Marilyn Bell Park Tennis Court	Casual Member
666	Navy Wharf Ct. / Bremner Blvd.	Annual Member
453	Front St E / Bayview Avenue	Annual Member
951	Marilyn Bell Park Tennis Court	Casual Member
1130	Ontario Place Blvd / Lake Shore Blvd W (East)	Casual Member
313	Palmerston Ave / Dundas St W	Annual Member

The key metrics of ride quantity, trip duration, and start station density were summarised and segmented by user type in tables 2, 3, and 4 respectively. Although not statistically tested, the difference between segments was large enough to warrant further exploration.

Table 2: Total number of rides by user type

User Type	Total number of rides
Casual Member	1032580
Annual Member	1499361

Table 3: Total number of rides by user type

User Type	Average duration (minutes)
Casual Member	15.58576
Annual Member	11.94579

Table 4: Top 10 most annual-dense start stations

Start Station	Number of casual rides	Number of annual rides
Westmount Ave / St Clair Ave W - SMART	41	226
Gailbraith Rd / KingGÇÖs College Cr. (U of T)	510	2055
Alvin Ave / St Clair Ave E - SMART	221	875
Simcoe St / Michael Sweet Ave - SMART	386	1517
Elm St/ University Ave (East)	530	2024
Belmont St / Davenport Rd - SMART	98	362
20 Charles St E - SMART	1031	3808
Baldwin St / Henry St - SMART	659	2400
Dentonia Park	77	256
Berkeley St / Dundas St E - SMART	601	1945

3 Discussion

3.1 Number of rides

While annual riders account for more overall rides than casual riders, annual-rider activity appears to peak before casual riders as seen in Figure 1. This was an unexpected finding; I assumed that annual riders would be more likely to have Bike Share woven into their routines, making them less responsive to factors like colder temperatures. Given the steep drop off in annual rides around early October, there may have been external events within the city that affected this segment more than casual riders; changes in return-to-office policies come to mind. Viewing historical data may provide another perspective on this specific point.

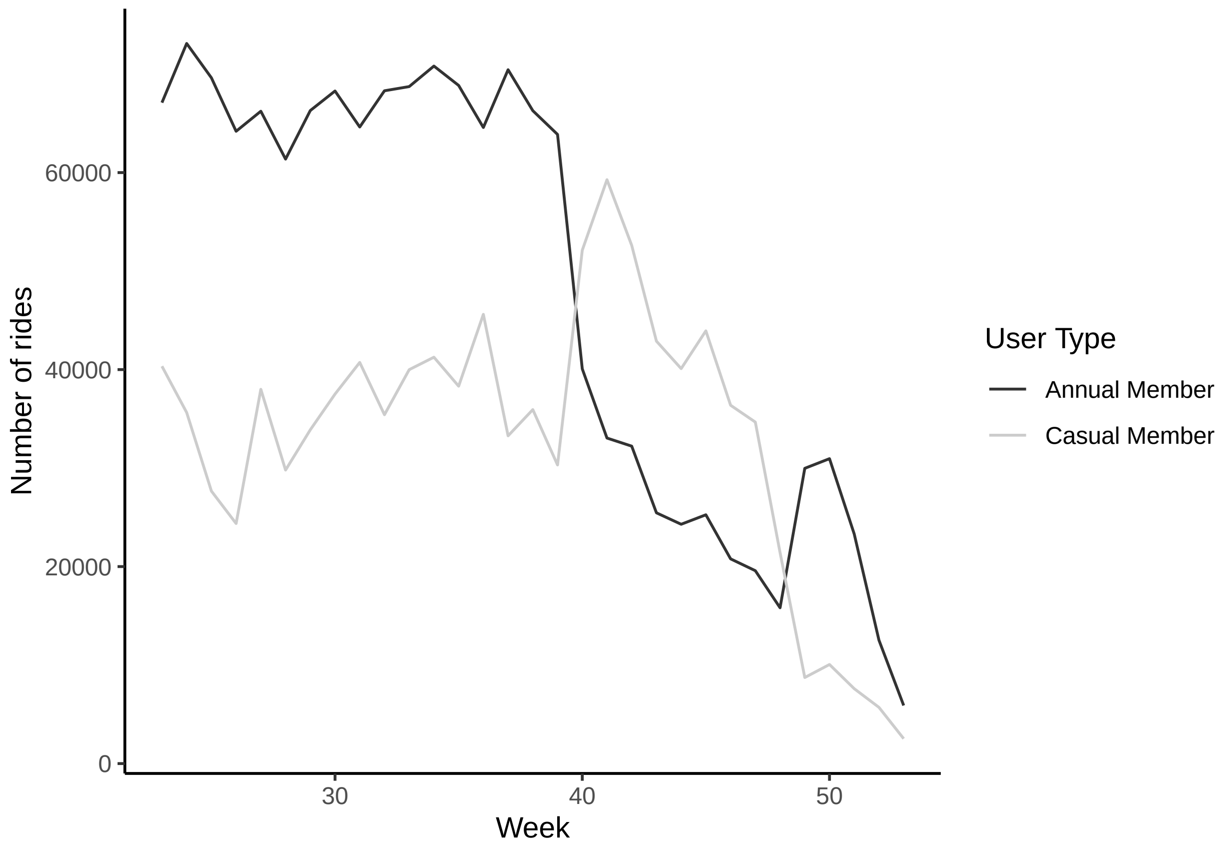

Figure 1: Total number of rides by user type

3.2 Trip Duration

I was curious to explore trip duration in depth, but realised that Bike Share's time limits - where users must dock their bikes within either 30 or 45 minutes of their start time to prevent accruing additional charges - would heavily skew results towards shorter durations. In a potentially confusing move, a 45-minute option was added to the historical 30-minute limit for some memberships in June of 2021, right at the beginning of this dataset (Sumi 2021). Rider behaviour may therefore still reflect the old half-hour limit; I was not aware of this change before researching it specifically.

Whilst most bike commutes will be under this limit, this rule masks the true duration of casual trips; for example, a casual user may take ~1.5 hours to bike along the waterfront, but will dock and undock their bike

2 additional times within that time window. This shows up as 3 separate trips in the data and significantly reduces duration-based metrics, as the average trip duration for that user will be under 30 minutes when they used the bike for 3 times that in reality.

Figure 2 below shows a histogram of trip durations segmented by annual and casual users. While casual rides are longer on average, one can see an unnatural drop off in rides at the 30-minute mark; this implies that users are artificially cutting their rides short to dock their bikes, meaning that true trip durations could easily be longer.

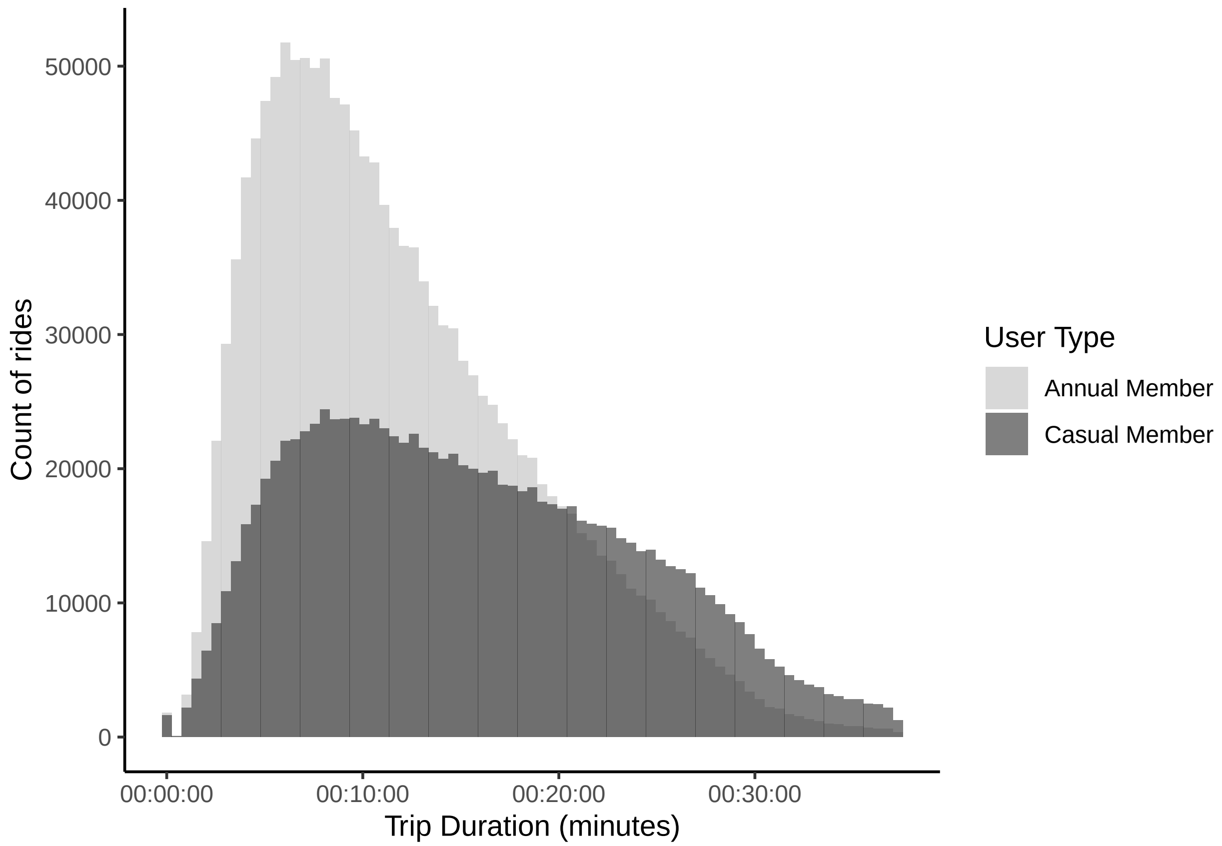

Figure 2: Histogram of ride duration by user type

This provides interesting implications when considering improving the Bike Share experience for casual users. Having to cut a ride short to find a station, dock, re-enter payment information, undock, before being able to continue is undoubtedly unpleasant - something that I have also experienced first hand (Elliot 2020). It might even discourage riders from fully utilising Toronto's extensive bikeshare network. Could there be a "casual pass" for these riders, or an extension to a 1 hour time limit? This would tackle a key pain point of the casual experience, perhaps encouraging riders to engage with an annual pass.

Another interesting trend was the decrease in trip duration as the year progressed, demonstrated below in Figure 3. This behaviour is expected, but casual riders exhibit a much steeper drop off. This is partly due to the higher average duration of casual riders, but also reveals the elasticity of casual demand based on what are likely weather-related factors. Note the steep climb in average casual rider duration towards the very end of the year. This may be a 2021-specific trend, so examining data from previous years would be helpful, but it suggests increased activity around the holiday period. Perhaps this natural demand can be encouraged by Bike Share with holiday-specific promotions, or more attention can be placed on ensuring that bike lanes around popular casual areas remain accessible and clear of snow.

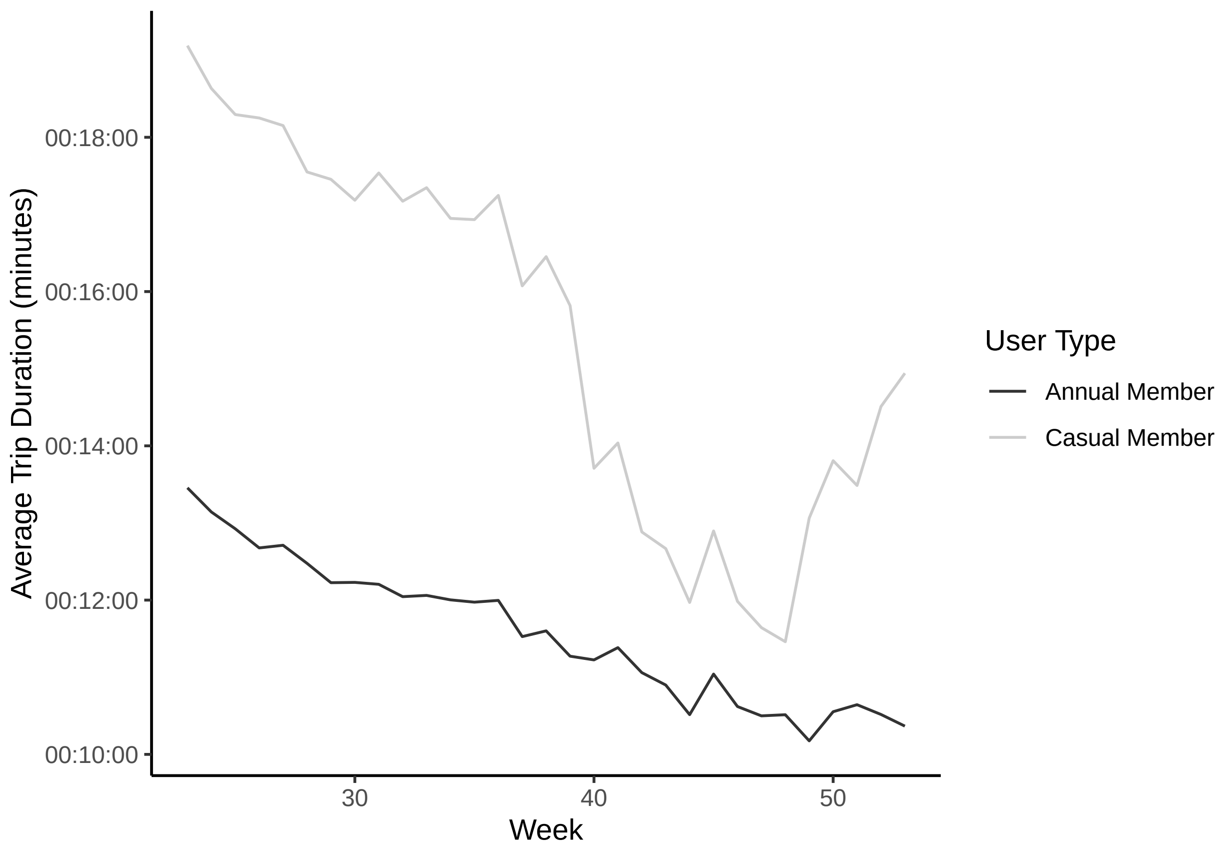

Figure 3: Average trip duration by user type

3.3 Location Hotspots

8 of the top 10 most popular bike stations are along the waterfront, where an extensive bike trail has meant safe and pleasant riding for regular and casual riders alike - since even before this century (Koropeski, Loane, and Gough 1998). Segmenting station popularity by user type reveals a network dominated by annual riders, adding detail to how Bike Share is being utilised around the city. The top 10 most trafficked station-user type pairs are split evenly between annual and casual riders, but the next 40 are dominated by annual riders. Ranking starting stations by the ratio of casual:annual riders then reveals that certain areas are clearly "annual-only" or "casual-only". This is demonstrated by figures 4 and 5, showing the top 20 stations for annual and casual ridership density respectively (density calculated as the ratio between user types). Both maps are produced courtesy of `OpenStreetMap` (Contributors 2017).

Figure 4: Top 20 start stations by annual rider density

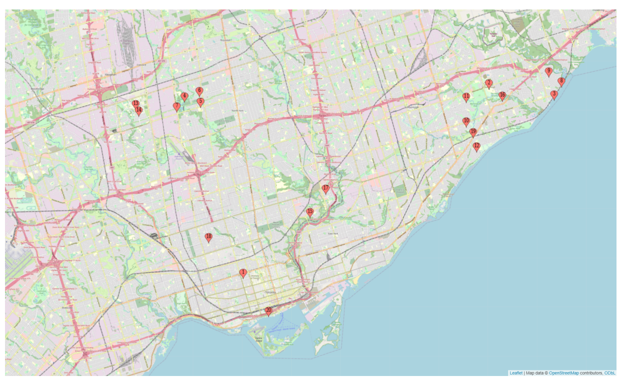

Figure 5: Top 20 start stations by casual rider density

This stark difference raises several possible areas of exploration. Above all, I am interested in knowing whether or not this network duality is intentional. I feel that Bike Share must acknowledge these dynamics as natural behaviour, and I would therefore like to dive deeper into how these two segments can be better served in their respective regions.

Take the station at King's College Crescent on our own UofT campus; given that just 20% of rides originating here are by casual riders, are there commuter-friendly actions that can be implemented? Apart from introducing potentially exciting features, such as lockers, simple actions such as ensuring that sufficient bikes are available at this station during peak hours can greatly aid annual riders who rely on Bike Share for commuting.

4 Conclusions

There is a clear distinction between how annual and casual riders use Bike Share, but both segments are largely served the same experience. As Toronto emerges from pandemic lockdowns and strives for a sustainable future, clean urban mobility solutions like Bike Share can be refined and improved to encourage greater adoption. Immediate findings from this analysis point to a restrictive and possibly confusing single-ride time limit, where riders must dock within 30 or 45 minutes of starting a specific segment. "Casual-only" or "annual-only" hotspots are also revealed, painting a clear picture of who is using the network at what location and spurring thoughts about what their needs could be.

Going forward, it is important to examine historical data beyond 2021; this will help determine whether the trends in this dataset are annual occurrences or a COVID-specific blip. The pattern of increasing ride duration at the very end of 2021 from casual riders is one potential anomaly, for example.

References

Contributors, OpenStreetMap. 2017. "Planet Dump Retrieved from Https://Planet.osm.org." https://www.openstreetmap.org.

Dia, Hussein, Michael Taylor, John Stone, Sekhar Somenahalli, and Stephen Cook. 2019. *Low Carbon Urban Mobility*. https://link.springer.com/chapter/10.1007/978-981-13-7940-6_14.

Dowle, Matt, Arun Srinivasan, Jan Gorecki, Michael Chirico, Pasha Stetsenko, Tom Short, Steve Lianoglou, et al. 2021. *Data.table: Extension of 'Data.frame'*. https://cran.r-project.org/package=data.table.

Elliot, Matt. 2020. "Bike Share Toronto Is Finally Thriving (and Those Electric Bikes Are Awesome, Too)." *Toronto Star*. https://www.thestar.com/opinion/contributors/2020/08/24/bike-share-toronto-is-finally-thriving-and-those-electric-bikes-are-awesome-too.html.

Firke, Sam, Bill Denney, Chris Haid, Ryan Knight, Malte Grosser, and Jonathan Zadra. 2021. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://cran.r-project.org/package=janitor.

Gelfand, Sharla. 2020. *Opendatatoronto: Access the City of Toronto Open Data Portal*. https://cran.r-project.org/package=opendatatoronto.

Hosford, Kate, and Meghan Winters. 2018. *Who Are Public Bicycle Share Programs Serving? An Evaluation of the Equity of Spatial Access to Bicycle Share Service Areas in Canadian Cities. Transportation Research Record: Journal of the Transportation Research Board*. https://journals.sagepub.com/doi/full/10.1177/0361198118783107.

Koropeski, Andrew, Gregg Loane, and Jim Gough. 1998. "Toronto - Making Room for All Modes." *Institute of Transportation Engineers. ITE Journal* 68 (6): 42.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Schwanen, Tim. 2021. *Achieving Just Transitions to Low-Carbon Urban Mobility. Nature Energy*. https://www.nature.com/articles/s41560-021-00856-z.

Sumi, Glenn. 2021. "Time's up for Bike Share Toronto Riders." *NowToronto*. https://nowtoronto.com/news/times-up-for-bike-share-toronto-riders.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2021. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Zhu, Hao, Thomas Travison, Timothy Tsai, Will Beasley, Yihui Xie, GuangChuang Yu, Stéphane Laurent, et al. 2021. *KableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://CRAN.R-project.org/package=dplyr.