

Low Income, Low Tree Cover? An Analysis on Neighbourhood Tree Cover and Median Income Profile

Asel Kushkeyeva and Shabrina J. Mardevi

9 February 2020

Abstract

The purpose of this project is to explore on Toronto neighbourhood's urban forestry status, through the variables of tree coverage and net median income profile. Our research question lies on whether neighbourhoods with lower median income status would have less aerial tree cover. To do this, we downloaded and analysed two 2011 datasets from Open Data Toronto; namely, (1) Wellbeing - Environment (Air Quality) and (2) Demographics - National Household Survey (NHS) Indicators. After conducting summary statistics and generating scatterplot graphs, we find a pattern in which lower numbers of tree coverage is found concentrated in lower end of median income neighbourhoods (~\$31,000-55,000), and that there is no tree cover below average in higher income neighbourhood with net median income of more than ~\$75,000. Nevertheless, we postulate that despite extremities, there needs to be more considerations and discussions in measuring relative greenery in Toronto neighbourhoods.

Introduction

City trees are essential for overall urban and residential qualities of life. They play a pivotal role, not only for health and environmental benefits such as improving air and water qualities. They are also important for community values, such as in promoting active lifestyle, making available of outdoor recreation spaces, and creating stronger sense of community (Harlan et al., 2006).

City of Toronto acknowledges the importance of protecting and expanding its urban forest. The city offers various tree planting grants and programs. In Treepedia project, measuring land covered by trees in 12 cities around the world, Toronto ranked fifth (CBC, 2017). Nevertheless, there are still further improvements to be made for tree/canopy distributions. Currently, Toronto aims to increase its canopy cover from 27% to approximately 30 to 40% of total land area (City of Toronto et al., 2013). It was also found that majority of city's canopy is on private lands (Tran, 2019). City of Toronto (2013) further reports that highest tree coverage are found in neighbourhoods with some of the highest average real estate values. This begs the question whether there is a distribution inequality in canopy cover in Toronto based neighbourhood income profile.

With that, in this project, we seek to define the role of neighbourhood socioeconomic status in determining their corresponding aerial tree cover. Our analysis, focusing on median after-tax income of different Toronto neighbourhoods and tree cover, revealed some noteworthy patterns. We hypothesized that people with higher after-tax median income lived in greener neighbourhoods. We observed a strong concentration in lower household median income bracket. Overall, the levels of income are not the sole determinants of larger tree coverage in neighbourhoods people live in.

Dataset

To address our research question on neighbourhood tree cover and socioeconomic profile, we used and merged two different datasets from Open Data Toronto catalogue.

Wellbeing - Environment (Air Quality)

Toronto Wellbeing - Environment dataset is a part of Open Data Toronto Wellbeing series. This dataset specifically presents data on air quality. According to Open Data Toronto, the dataset was last refreshed in December 31st, 2014. The data was made available by Toronto Public Health, Toronto Parks, Forestry, and Recreation, Federal Ministry of the Environment, and Toronto Water.

This dataset consists of 3 sheets, namely; Indicator Metadata, raw data of 2008, and raw data of 2011. The Indicator Metadata contains detailed information for each column in the dataset, as well as where we retrieved the unit measure for Tree Cover. We selected the 2011 as it is the most recent period available. In total, there are 140 rows, representing Neighbourhood as main categories, and 8 columns for air quality variables, such as those relating to green spaces, pollution emissions, and tree cover. For the purpose of this project, we used the column/variable Tree Cover in square metres (sq.m) for canopy cover measures.

Demographics - National Household Survey (NHS) Indicators

Toronto Wellbeing - NHS Demographics is also a part of Open Data Toronto Wellbeing series. This dataset is a collection of voluntary 2011 National Household Survey (NHS) responses, conducted by Statistics Canada. According to Open Data Toronto, the dataset was last refreshed in April 30th, 2016.

Similar to previous dataset, there are in total of 140 rows representing Neighbourhood as main categories. Nonetheless, there are 48 columns for demographics. The types of demographic information available in this dataset are such as total population, age, ethnicity, education, and income. For this project, we selected median after-tax household income variable for representation of socioeconomic profile in different neighbourhoods.

Analysis

Both of the variables that we are to explore are numeric/continuous. After executing summary statistics of both variables for later Exploratory Data Analysis benchmarking, we found several notable results and skewness of the values of data. In the Exploratory Data Analysis section, it will be shown how the following findings may affect the analysis.

In Table 1. it is shown that the net median income gap in Toronto is considerably wide, as the minimum is 30,794 Canadian Dollars (CAD) and the maximum is 161,448 CAD. This is further shown in Table 1. by how the 3rd Quartile of net median household income at 59,963 CAD is less than half of the maximum net median income at 161,448 CAD.

Table 1. Summary Statistics of \$ Net (After-Tax) Median Income Household

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	30794	46690	52660	55426	59963	161448

In Table 2., similar cases of extreme results were also found. The range of minimum and maximum aerial tree covers are extremely wide, with 61,616 square metres (sq.m) at minimum and 12,888,044 sq.m. at maximum. The differences between minimum value at 61,616 sq.m and lower quartile at 523,162 sq.m, and

between upper quartile at 1,686,216 sq.m and maximum at 12,888,044 sq.m, further show the extremes and unequal distributions of data.

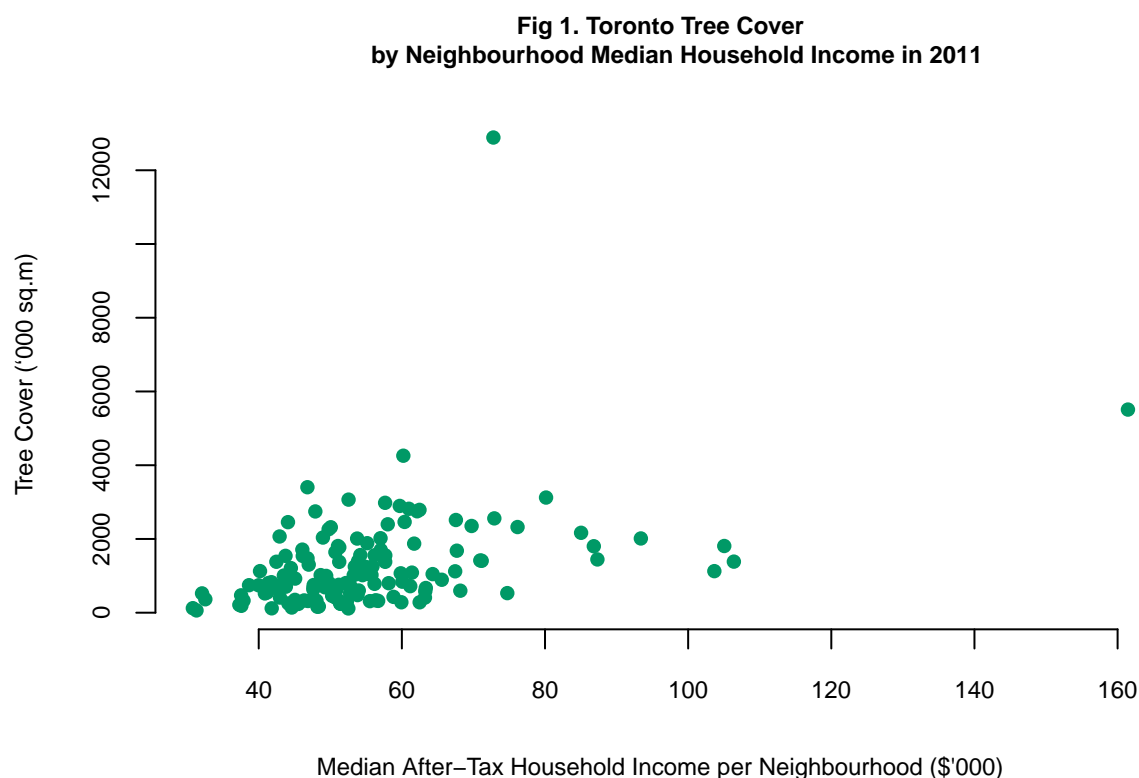
Table 2. Summary Statistics of Tree Covers in square metres

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	61616	523162	1017744	1281438	1686216	12888044

Exploratory Data Analysis

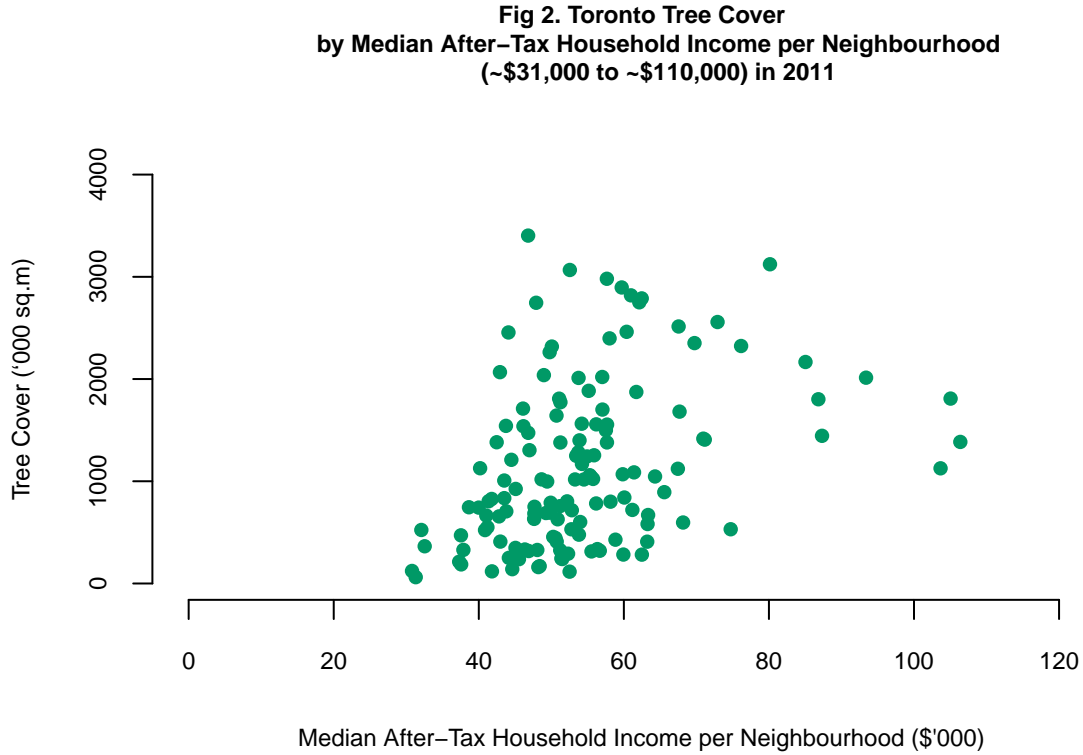
Figure 1. was executed to see whether there is some sort of relationship/correlation between Median After-Tax Household Income and Tree Cover in Toronto neighbourhoods. As shown in Figure 1., there is a slight general increasing trend, suggesting that there is a relationship between neighbourhood net median income profile and canopy cover. However, this does not seem to be a strong correlation as there are some cluster of plots of high aerial tree cover (approximately more than the 3rd Quartile value of 1,686,216 sq.m, refer to Table 2.) in neighbourhoods with low to middle net median income approximately between 46,690 CAD (1st Quartile, Table 1.) to 52,660 CAD (median, Table 1.).

Figure 1. Toronto Tree Cover by Neighbourhood Median Household Income in 2011



Nevertheless, there are still noteworthy findings in the graph relating to the main hypothesis. Since all of the lower than average aerial tree cover (below 1,281,438 sq.m, Table 2.) are approximately concentrated in neighbourhoods with lower to middle net median income, Figure 2. is executed with emphasis on the indicated income bracket. In neighbourhoods with above approximately \$75,000 net median income, Figure 1. and 2. also show that there are exclusively aerial tree covers that are above average.

Figure 2. Toronto Tree Cover by Neighbourhood Median Household Income (~\$31,000 to ~\$110,000) in 2011



Discussion and Limitations

Despite seemingly weak and/or inconclusive correlations, the exclusivities of values at the extreme ends of the graph present new points of discussions. In which, inequality of tree distributions perhaps does exist but needs further explorations of other variables. In future studies, further aspects of urban planning and social factors should be considered. For example, neighbourhood relativity to total area or aerial greenery should be measured. Moreover, Harlan et al. (2006) mentioned that not only lower socioeconomic groups were more likely living in less canopy-covered areas, but also those of ethnic minorities. Therefore, as for social factors, ethnic diversity and/or visible minorities in the neighbourhood can possibly be considered as variables.

These suggestions also relate to the limitations in this project. Aerial measure of canopy cover should be relative to total area and urban setting of neighbourhoods, instead of solely by the metric values. For instance, some neighbourhoods may be multiple times the size of others, and some neighbourhoods may be located closer to the inner cities where there is less greenery as opposed to the suburbs. Another possible factor/considerable variable and current limitations are the responsiveness and participation rates of residents in tree-planting efforts. This was supported as Tran (2019) states that a common characteristics of tree-planting initiatives should start from home and business owners in neighbourhoods.

Appendix A: Code

Setup and Preparation

```
knitr::opts_chunk$set(echo = TRUE)

my_packages <- c("opendatatoronto", "dplyr", "tidyr")
for (pkg in my_packages) {if (!pkg %in% installed.packages()) {install.packages(pkg)}}

library(opendatatoronto)
library(dplyr)
library(tidyr)
```

Dataset Download and Merging

```
demographics_packages <- search_packages("demographics")
demographics_packages
```

```
demographics_nhs_package<-search_packages("Wellbeing Toronto - Demographics: NHS Indicators")
demographics_nhs_resources<-demographics_nhs_package %>% list_package_resources()
demographics_nhs_resources
```

```
demographics_nhs_dataset<-get_resource("8d838d5c-20da-48bd-b4a7-3e66b1c25b55")
demographics_nhs_dataset = demographics_nhs_dataset[-1,]
head(demographics_nhs_dataset)
```

```
demographics_nhs_dataset<-get_resource("8d838d5c-20da-48bd-b4a7-3e66b1c25b55")
head(demographics_nhs_dataset)
```

```
environment_packages <- search_packages("environment")

environment_wellbeing <- environment_packages %>%
  list_package_resources() %>%
  get_resource()

enviro_wellbeing2011 <- environment_wellbeing[["RawData-Ref Period 2011"]]
names(enviro_wellbeing2011)[names(enviro_wellbeing2011) == 'Neighbourhood Id'] <- 'Hood#'
head(enviro_wellbeing2011)
```

Cleaning

Sorting by Neighbourhood Number

```
new_data<-merge(demographics_nhs_dataset, enviro_wellbeing2011, by = "Hood#")
new_data$`Hood#`<-as.integer(new_data$`Hood#`)
new_data <- new_data[order(new_data$`Hood#`),]
```

Remove Additional Column for Labels

```
demographics_nhs_dataset = demographics_nhs_dataset[-1,]
```

Changing Class Variables from Factor to Numeric

```
new_data$`Tree Cover`<-as.numeric(new_data$`Tree Cover`)  
new_data$`Median after-tax household income `$<-as.numeric(new_data$`Median after-tax household income $`)
```

Summary Statistic Tables

Table 1. Summary Statistics of \$ Net (After-Tax) Median Income Household

```
summary(new_data$`Median after-tax household income `$)
```

Table 2. Summary Statistics of Tree Covers in square metres

```
summary(new_data$`Tree Cover`)
```

Appendix B: Graphs and their Codes

Figure 1. Toronto Tree Cover by Neighbourhood Median Household Income in 2011

```
x <- new_data$`Median after-tax household income $`/1000  
y <- new_data$`Tree Cover`/1000
```

```
plot(x, y, main = "Fig 1. Toronto Tree Cover  
by Neighbourhood Median Household Income in 2011",  
xlab = "Median After-Tax Household Income per Neighbourhood ($'000)",  
ylab = "Tree Cover (`000 sq.m)", pch = 16, frame = FALSE, col = "#009966",  
cex.main=0.75, cex.lab=0.75, cex.axis=0.75)
```

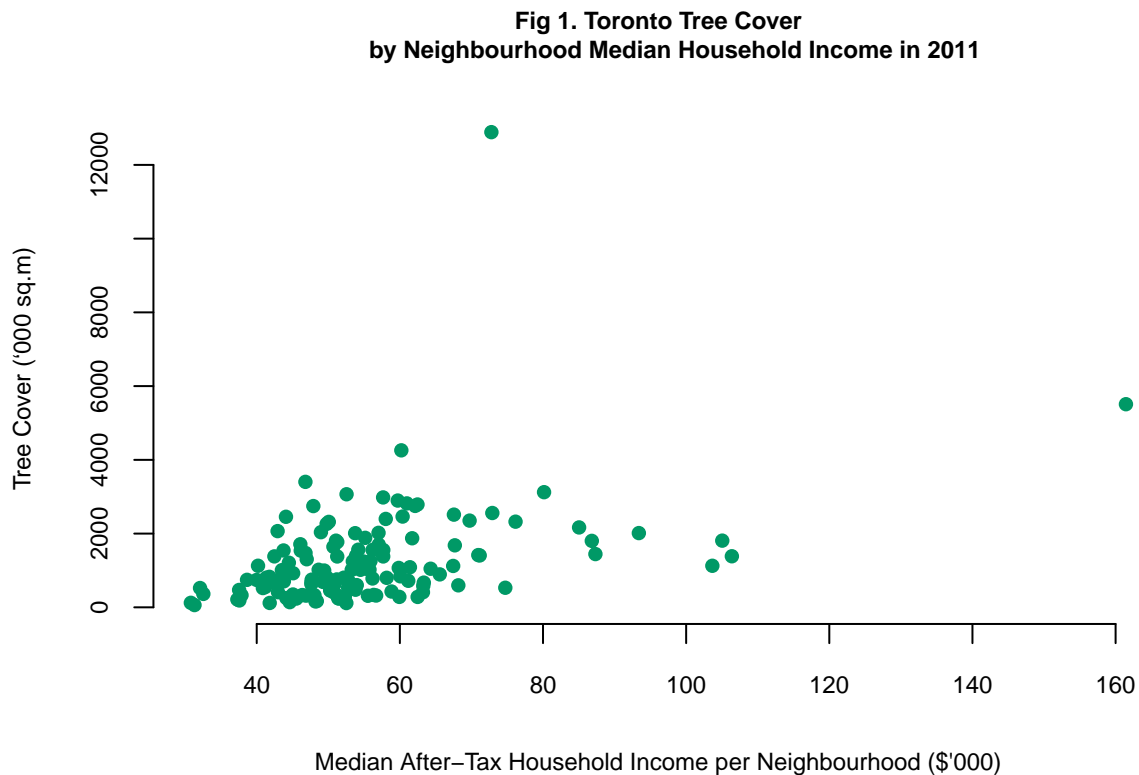
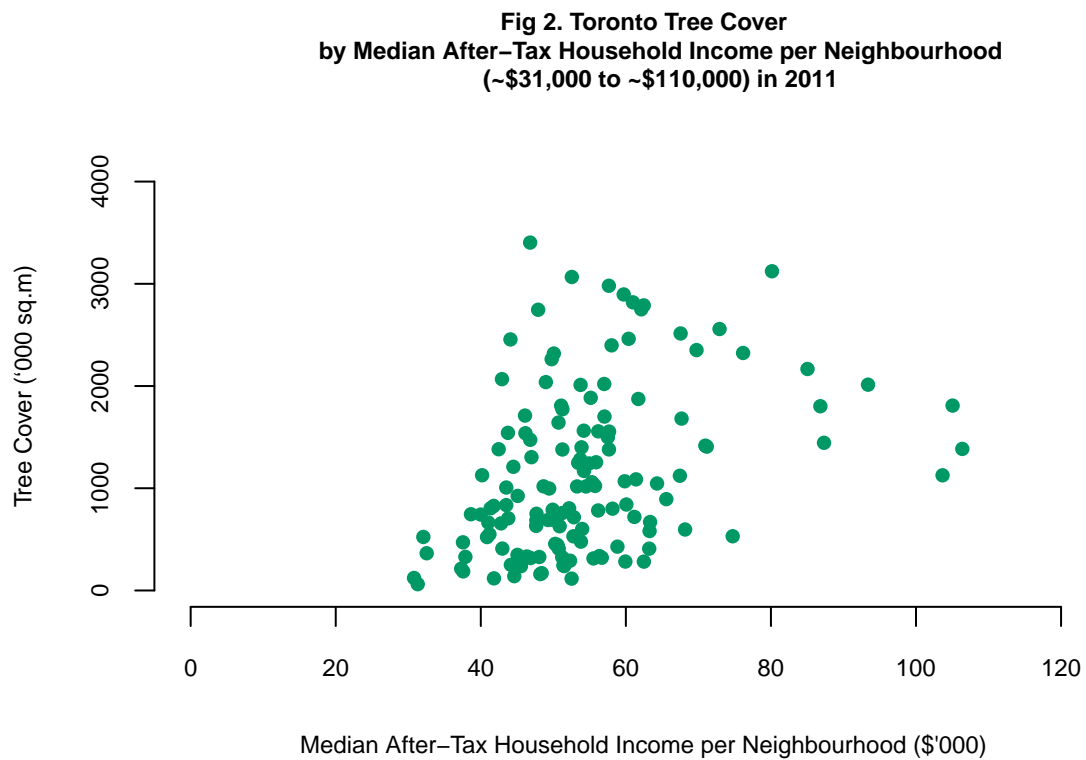


Figure 2. Toronto Tree Cover by Neighbourhood Median Household Income (~\$31,000 to ~\$110,000) in 2011

```
par(mar=c(5,6,6,1))
plot(x, y, main = "Fig 2. Toronto Tree Cover
  by Median After-Tax Household Income per Neighbourhood
  (~$31,000 to ~$110,000) in 2011",
  xlab = "Median After-Tax Household Income per Neighbourhood ($'000)",
  ylab = "Tree Cover (`000 sq.m)", pch = 16, frame = FALSE,
  xlim=c(0,125), ylim=c(0,4000), col="#009966",
  cex.main=0.75, cex.lab=0.75, cex.axis=0.75)
```



References

- City of Toronto, Parks, Forestry and Recreation, Urban Forestry. Toronto, Ontario. (2013). Every Tree Counts: A Portrait of Toronto’s Urban Forest. <https://www.toronto.ca/wp-content/uploads/2017/12/92de-every-tree-counts-portrait-of-torontos-urban-forest.pdf>
- City News. (2019, August 19). Toronto and feds pledge \$30M to rehabilitate city’s shorelines, tree canopy. <https://toronto.citynews.ca/2019/08/19/toronto-and-feds-pledge-30m-to-rehabilitate-citys-shorelines-tree-canopy/>
- DaSilva, L. (2017, January 4). Toronto beats NYC, Paris, London in new tree ranking, but experts say more work needed. CBC News. <https://www.cbc.ca/news/canada/toronto-tree-canopy-ranked-fifth-treepedia-project-1.3921714>
- Gelfand, Sh. (2019). opendatatoronto: Access the City of Toronto Open Data Portal. R package version 0.1.1. <https://CRAN.R-project.org/package=opendatatoronto>
- Harlan, S. L., Brazel, A. J., Prashad, L., Stefanov, W. L., & Larsen, L. (2006). Neighbourhood microclimates and vulnerability to heat stress. *Social Science & Medicine*, 63, 2847-2863
- R Core Team. (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.
- Toronto Urban Forestry. City of Toronto. (2019). Tree Planting Strategy. <https://www.toronto.ca/business-economy/partnerships-sponsorships-donations/partner-2/parks-environment/tree-planting-strategy/>
- Tran, T. (2019). Planting trees in low-income communities in Toronto: A look at the pilot TCHC Planting and Stewardship Initiative. http://www.smithlab.ca/uploads/1/2/0/5/120564888/truc-lam_tran_capstone_2018.pdf
- Wickham et al. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43), 1686, <https://doi.org/10.21105/joss.01686>
- Wickham, H., François, R., Henry, L. and Müller, K. (2019). dplyr: A Grammar of Data Manipulation. R package version 0.8.3. <https://CRAN.R-project.org/package=dplyr>
- Wickham, H. and Henry, L. (2019). tidyr: Tidy Messy Data. R package version 1.0.0. <https://CRAN.R-project.org/package=tidyr>
- Xie, Y. (2014) knitr: A Comprehensive Tool for Reproducible Research in R. In Victoria Stodden, Friedrich Leisch and Roger D. Peng, editors, *Implementing Reproducible Computational Research*. Chapman and Hall/CRC. ISBN 978-1466561595
- Xie, Y. (2015) *Dynamic Documents with R and knitr*. 2nd edition. Chapman and Hall/CRC. ISBN 978-1498716963
- Xie, Y. (2019). knitr: A General-Purpose Package for Dynamic Report Generation in R. R package version 1.26.