Stochastic Topographical Mobility Analysis (STMA) in the Copenhagen Metropolitan Area

Matthew Papesh

2025-04-23

#Include needed libraries below:

suppressWarnings({  
 if(!require("ggplot2")) {  
 install.packages("ggplot2", repo="http://cran.us.r-project.org", quite=TRUE, message=FALSE)  
 require("ggplot2")  
 }  
 if(!require("ggtext")) {  
 install.packages("ggtext", repo="http://cran.us.r-project.org", quite=TRUE, message=FALSE)  
 require("ggtext")  
 }  
 if(!require("tidyverse")) {  
 install.packages("tidyverse", repo="http://cran.us.r-project.org", quite=TRUE, message=FALSE)  
 require("tidyverse")  
 }  
 if(!require("dplyr")) {   
 install.packages("dplyr", repo="http://cran.us.r-project.org", quite=TRUE, message=FALSE)  
 require("dplyr")  
 }  
 if(!require("DescTools")) {  
 install.packages("DescTools", repo="http://cran.us.r-project.org", quite=TRUE, message=FALSE)  
 require("DescTools")  
 }  
 if(!require("gridExtra")) {  
 install.packages("gridExtra", repo="http://cran.us.r-project.org", quite=TRUE, message=FALSE)  
 require("gridExtra")  
 }  
 if(!require("knitr")) {  
 install.packages("knitr", repo="http://cran.us.r-project.org", quite=TRUE, message=FALSE)  
 require("knitr")  
 }  
})

## Loading required package: ggplot2

## Loading required package: ggtext

## Loading required package: tidyverse

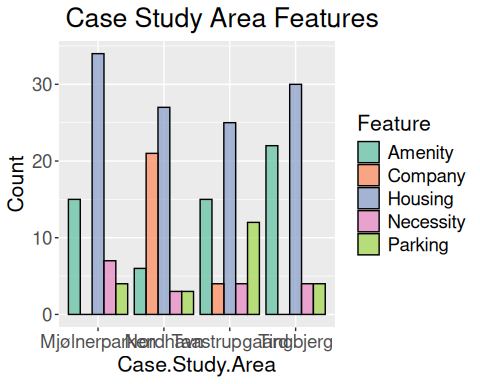
## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──  
## ✔ dplyr 1.1.4 ✔ readr 2.1.5  
## ✔ forcats 1.0.0 ✔ stringr 1.5.1  
## ✔ lubridate 1.9.4 ✔ tibble 3.2.1  
## ✔ purrr 1.0.4 ✔ tidyr 1.3.1  
## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()  
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors  
## Loading required package: DescTools  
##   
## Loading required package: gridExtra  
##   
##   
## Attaching package: 'gridExtra'  
##   
##   
## The following object is masked from 'package:dplyr':  
##   
## combine  
##   
##   
## Loading required package: knitr

# Import mobility data:

# imported node data by case study area   
nordhavn\_nodes = read.csv("./clean\_network\_mobility\_data/nordhavn\_nodes.csv")  
mjolnerparken\_nodes = read.csv("./clean\_network\_mobility\_data/mjolnerparken\_nodes.csv")  
tingbjerg\_nodes = read.csv("./clean\_network\_mobility\_data/tingbjerg\_nodes.csv")  
taastrupgaard\_nodes = read.csv("./clean\_network\_mobility\_data/taastrupgaard\_nodes.csv")  
# imported trip data by case study area   
nordhavn\_trips = read.csv("./stma\_results/nordhavn\_results.csv")  
mjolnerparken\_trips = read.csv("./stma\_results/mjolnerparken\_results.csv")  
tingbjerg\_trips = read.csv("./stma\_results/tingbjerg\_results.csv")  
taastrupgaard\_trips = read.csv("./stma\_results/taastrupgaard\_results.csv")

# qaulitative location data:

# count the number of features of a value type in a given vector  
count\_feat <- function(vector, type\_value) {  
 # vector : vector name   
 # type\_value : the value to count inside given vector  
 sum(vector == type\_value, na.rm=TRUE)  
}  
  
# vectors of feature types for each case study  
nord\_f <- (nordhavn\_nodes %>% filter(node\_id == "inner" | node\_id == "frontier"))$type  
mjol\_f <- (mjolnerparken\_nodes %>% filter(node\_id == "inner" | node\_id == "frontier"))$type  
ting\_f <- (tingbjerg\_nodes %>% filter(node\_id == "inner" | node\_id == "frontier"))$type  
taast\_f <- (taastrupgaard\_nodes %>% filter(node\_id == "inner" | node\_id == "frontier"))$type  
  
data <- data.frame(  
 Case.Study.Area = rep(c("Nordhavn","Mjølnerparken","Tingbjerg","Taastrupgaard"), each = 5),  
 Feature = rep(c("Housing","Necessity","Amenity","Company","Parking"), 4),  
 Count = c(  
 count\_feat(nord\_f, "house"), count\_feat(nord\_f, "necessity"),  
 count\_feat(nord\_f, "amenity"), count\_feat(nord\_f, "company"),  
 count\_feat(nord\_f, "parking"),  
  
 count\_feat(mjol\_f, "house"), count\_feat(mjol\_f, "necessity"),  
 count\_feat(mjol\_f, "amenity"), count\_feat(mjol\_f, "company"),  
 count\_feat(mjol\_f, "parking"),  
  
 count\_feat(ting\_f, "house"), count\_feat(ting\_f, "necessity"),  
 count\_feat(ting\_f, "amenity"), count\_feat(ting\_f, "company"),  
 count\_feat(ting\_f, "parking"),  
  
 count\_feat(taast\_f, "house"), count\_feat(taast\_f, "necessity"),  
 count\_feat(taast\_f, "amenity"), count\_feat(taast\_f, "company"),  
 count\_feat(taast\_f, "parking")  
 )  
)  
  
ggplot(data, aes(x = Case.Study.Area, y = Count, fill = Feature)) +  
 labs(title = " Case Study Area Features") +  
 geom\_col(position = "dodge", colour = "black", alpha = 0.75) +  
 scale\_fill\_brewer(palette = "Set2") +   
 scale\_color\_brewer(palette = "Set2") +   
 theme(  
 plot.title = element\_text(size = 20), # Title text size  
 axis.title.x = element\_text(size = 16), # X-axis label size  
 axis.title.y = element\_text(size = 16), # Y-axis label size  
 axis.text.x = element\_text(size = 14), # X-axis tick labels size  
 axis.text.y = element\_text(size = 14), # Y-axis tick labels size  
 legend.title = element\_text(size = 16), # Legend title size  
 legend.text = element\_text(size = 14) # Legend item labels size  
)

 #

# case study area data grouped  
areas <- data\_frame(  
 distance\_km = c(nordhavn\_trips$distance\_km, mjolnerparken\_trips$distance\_km, tingbjerg\_trips$distance\_km, taastrupgaard\_trips$distance\_km),  
 time\_min = c(nordhavn\_trips$time\_min, mjolnerparken\_trips$time\_min, tingbjerg\_trips$time\_min, taastrupgaard\_trips$time\_min),  
 trip\_speed = c(nordhavn\_trips$distance\_km/nordhavn\_trips$time\_min, mjolnerparken\_trips$distance\_km/mjolnerparken\_trips$time\_min, tingbjerg\_trips$distance\_km/tingbjerg\_trips$time\_min, taastrupgaard\_trips$distance\_km/taastrupgaard\_trips$time\_min),  
 trip\_type = c(nordhavn\_trips$sample\_id, mjolnerparken\_trips$sample\_id, tingbjerg\_trips$sample\_id, taastrupgaard\_trips$sample\_id),  
 area\_name = c(rep("Nordhavn", nrow(nordhavn\_trips)), rep("Mjølnerparken", nrow(mjolnerparken\_trips)), rep("Tingbjerg", nrow(tingbjerg\_trips)), rep("Taastrupgaard", nrow(taastrupgaard\_trips)))  
)

## Warning: `data\_frame()` was deprecated in tibble 1.1.0.  
## ℹ Please use `tibble()` instead.  
## This warning is displayed once every 8 hours.  
## Call `lifecycle::last\_lifecycle\_warnings()` to see where this warning was  
## generated.

# case study area data filtered by trip  
walk\_areas <- areas %>% filter(trip\_type == "walk")  
bike\_areas <- areas %>% filter(trip\_type == "bike")  
local\_transit\_areas <- areas %>% filter(trip\_type == "local")  
commute\_areas <- areas %>% filter(trip\_type == "commute")  
alt\_commute\_areas <- areas %>% filter(trip\_type == "alt\_commute")  
# rename area for alternative commute   
alt\_commute\_areas$area\_name = paste0(alt\_commute\_areas$area\_name, rep(" by Train", nrow(alt\_commute\_areas)))  
# combine alternative commutes back into main list   
commute\_areas <- rbind(commute\_areas, alt\_commute\_areas)

# create density charts for test scenarios comparing case study areas

density\_chart\_walk <- ggplot(data = walk\_areas, aes(x = distance\_km, fill = area\_name, color = area\_name)) +  
 geom\_density(alpha = 0.15, size = 1) +   
 scale\_fill\_brewer(palette = "Set2") +   
 scale\_color\_brewer(palette = "Set2") +   
 theme\_minimal() +  
 labs(x = "Distance (km)", y = "Frequency") +  
 theme(  
 plot.title = element\_text(size = 20), # Title text size  
 axis.title.x = element\_text(size = 16), # X-axis label size  
 axis.title.y = element\_text(size = 16), # Y-axis label size  
 axis.text.x = element\_text(size = 14), # X-axis tick labels size  
 axis.text.y = element\_text(size = 14), # Y-axis tick labels size  
 legend.title = element\_text(size = 16), # Legend title size  
 legend.text = element\_text(size = 14) # Legend item labels size  
 )

## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.  
## ℹ Please use `linewidth` instead.  
## This warning is displayed once every 8 hours.  
## Call `lifecycle::last\_lifecycle\_warnings()` to see where this warning was  
## generated.

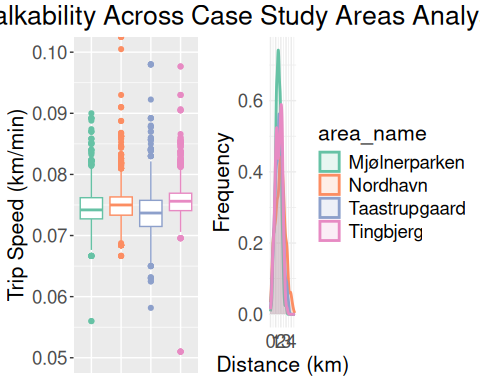
density\_chart\_bike <- ggplot(data = bike\_areas, aes(x = distance\_km, fill = area\_name, color = area\_name)) +  
 geom\_density(alpha = 0.15, size = 1) +   
 scale\_fill\_brewer(palette = "Set2") +   
 scale\_color\_brewer(palette = "Set2") +   
 theme\_minimal() +  
 labs(x = "Distance (km)", y = "Frequency") +   
 theme(  
 plot.title = element\_text(size = 20), # Title text size  
 axis.title.x = element\_text(size = 16), # X-axis label size  
 axis.title.y = element\_text(size = 16), # Y-axis label size  
 axis.text.x = element\_text(size = 14), # X-axis tick labels size  
 axis.text.y = element\_text(size = 14), # Y-axis tick labels size  
 legend.title = element\_text(size = 16), # Legend title size  
 legend.text = element\_text(size = 14) # Legend item labels size  
 )  
  
density\_chart\_local\_transit <- ggplot(data = local\_transit\_areas, aes(x = distance\_km, fill = area\_name, color = area\_name)) +  
 geom\_density(alpha = 0.15, size = 1) +   
 scale\_fill\_brewer(palette = "Set2") +   
 scale\_color\_brewer(palette = "Set2") +   
 theme\_minimal() +  
 labs(x = "Distance (km)", y = "Frequency") +   
 theme(  
 plot.title = element\_text(size = 20), # Title text size  
 axis.title.x = element\_text(size = 16), # X-axis label size  
 axis.title.y = element\_text(size = 16), # Y-axis label size  
 axis.text.x = element\_text(size = 14), # X-axis tick labels size  
 axis.text.y = element\_text(size = 14), # Y-axis tick labels size  
 legend.title = element\_text(size = 16), # Legend title size  
 legend.text = element\_text(size = 14) # Legend item labels size  
 )  
  
density\_chart\_commute <- ggplot(data = commute\_areas, aes(x = distance\_km, fill = area\_name, color = area\_name)) +  
 geom\_density(alpha = 0.15, size = 1) +   
 scale\_fill\_brewer(palette = "Set2") +   
 scale\_color\_brewer(palette = "Set2") +   
 theme\_minimal() +  
 labs(x = "Distance (km)", y = "Frequency") +   
 theme(  
 plot.title = element\_text(size = 20), # Title text size  
 axis.title.x = element\_text(size = 16), # X-axis label size  
 axis.title.y = element\_text(size = 16), # Y-axis label size  
 axis.text.x = element\_text(size = 14), # X-axis tick labels size  
 axis.text.y = element\_text(size = 14), # Y-axis tick labels size  
 legend.title = element\_text(size = 16), # Legend title size  
 legend.text = element\_text(size = 14) # Legend item labels size  
 )

# create boxplots for test scenarios comparing case study areas

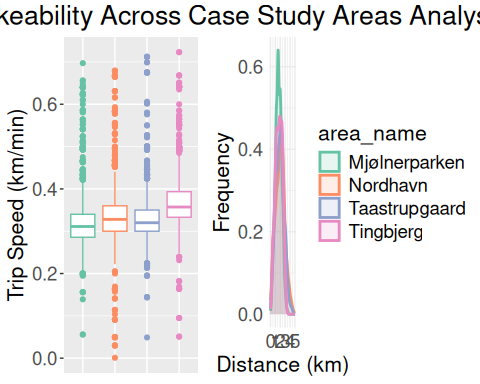
boxplot\_walk <- ggplot(data = walk\_areas, aes(x = area\_name, y = trip\_speed, color = area\_name)) +   
 geom\_boxplot() +   
 scale\_fill\_brewer(palette = "Set2") +   
 scale\_color\_brewer(palette = "Set2") +   
 coord\_cartesian(ylim = c(0.05, 0.10)) +  
 xlab("Case Study Area") +   
 ylab("Trip Speed (km/min)") +   
 theme(  
 legend.position = "none",   
 axis.title.x = element\_blank(),   
 axis.text.x = element\_blank(),   
 axis.ticks.x = element\_blank(),  
 axis.title.y = element\_text(size = 16), # Y-axis label size  
 axis.text.y = element\_text(size = 14), # Y-axis tick labels size  
 legend.title = element\_text(size = 16), # Legend title size  
 legend.text = element\_text(size = 14) # Legend item labels size  
 )  
  
boxplot\_bike <- ggplot(data = bike\_areas, aes(x = area\_name, y = trip\_speed, color = area\_name)) +   
 geom\_boxplot() +   
 scale\_fill\_brewer(palette = "Set2") +   
 scale\_color\_brewer(palette = "Set2") +   
 xlab("Case Study Area") +   
 ylab("Trip Speed (km/min)") +   
 theme(  
 legend.position = "none",   
 axis.title.x = element\_blank(),   
 axis.text.x = element\_blank(),   
 axis.ticks.x = element\_blank(),  
 axis.title.y = element\_text(size = 16), # Y-axis label size  
 axis.text.y = element\_text(size = 14), # Y-axis tick labels size  
 legend.title = element\_text(size = 16), # Legend title size  
 legend.text = element\_text(size = 14) # Legend item labels size  
 )  
  
boxplot\_local\_transit <- ggplot(data = local\_transit\_areas, aes(x = area\_name, y = trip\_speed, color = area\_name)) +   
 geom\_boxplot() +   
 scale\_fill\_brewer(palette = "Set2") +   
 scale\_color\_brewer(palette = "Set2") +   
 xlab("Case Study Area") +   
 ylab("Trip Speed (km/min)") +   
 theme(  
 legend.position = "none",   
 axis.title.x = element\_blank(),   
 axis.text.x = element\_blank(),   
 axis.ticks.x = element\_blank(),  
 axis.title.y = element\_text(size = 16), # Y-axis label size  
 axis.text.y = element\_text(size = 14), # Y-axis tick labels size  
 legend.title = element\_text(size = 16), # Legend title size  
 legend.text = element\_text(size = 14) # Legend item labels size  
 )  
  
boxplot\_commute <- ggplot(data = commute\_areas, aes(x = area\_name, y = trip\_speed, color = area\_name)) +   
 geom\_boxplot() +   
 scale\_fill\_brewer(palette = "Set2") +   
 scale\_color\_brewer(palette = "Set2") +   
 xlab("Case Study Area") +   
 ylab("Trip Speed (km/min)") +   
 theme(legend.position = "none", axis.title.x = element\_blank(), axis.text.x = element\_blank(), axis.ticks.x = element\_blank()) +   
 theme(  
 plot.title = element\_text(size = 20), # Title text size  
 axis.title.x = element\_text(size = 16), # X-axis label size  
 axis.title.y = element\_text(size = 16), # Y-axis label size  
 axis.text.x = element\_text(size = 14), # X-axis tick labels size  
 axis.text.y = element\_text(size = 14), # Y-axis tick labels size  
 legend.title = element\_text(size = 16), # Legend title size  
 legend.text = element\_text(size = 14) # Legend item labels size  
 )

# visualize STMA results:

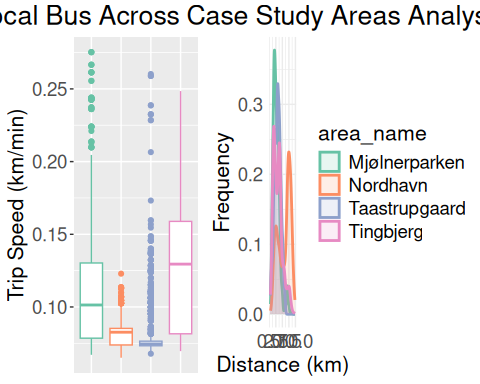
library(grid)  
  
walk\_title <- textGrob(  
 "Walkability Across Case Study Areas Analysis",  
 gp = gpar(fontsize = 20)  
)  
bike\_title <- textGrob(  
 "Bikeability Across Case Study Areas Analysis",   
 gp = gpar(fontsize = 20)  
)  
local\_transit\_title <- textGrob(  
 "Local Bus Across Case Study Areas Analysis",  
 gp = gpar(fontsize = 20)  
)  
city\_commute\_title <- textGrob(  
 "City Commute Across Case Study Areas Analysis",  
 gp = gpar(fontsize = 20)  
)  
  
grid.arrange(boxplot\_walk, density\_chart\_walk, ncol=2, widths=c(6,8), top=walk\_title)



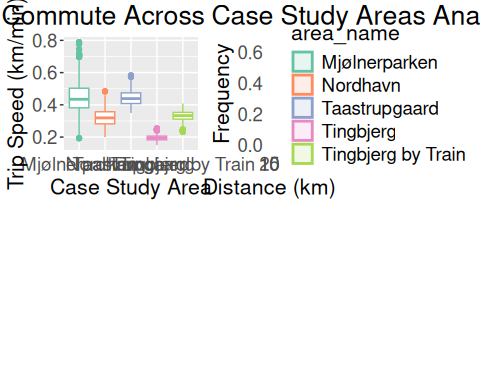
grid.arrange(boxplot\_bike, density\_chart\_bike, ncol=2, widths=c(6,8), top=bike\_title)



grid.arrange(boxplot\_local\_transit, density\_chart\_local\_transit, ncol=2, widths=c(6,8), top=local\_transit\_title)



grid.arrange(boxplot\_commute, density\_chart\_commute, nrow=2, widths=c(6,8), top=city\_commute\_title)

 # run kruskal wallis for statistical difference in trip efficiencies across case study areas

kruskal.test(data = walk\_areas, trip\_speed ~ area\_name)

##   
## Kruskal-Wallis rank sum test  
##   
## data: trip\_speed by area\_name  
## Kruskal-Wallis chi-squared = 545.92, df = 3, p-value < 2.2e-16

kruskal.test(data = bike\_areas, trip\_speed ~ area\_name)

##   
## Kruskal-Wallis rank sum test  
##   
## data: trip\_speed by area\_name  
## Kruskal-Wallis chi-squared = 1026, df = 3, p-value < 2.2e-16

kruskal.test(data = local\_transit\_areas, trip\_speed ~ area\_name)

##   
## Kruskal-Wallis rank sum test  
##   
## data: trip\_speed by area\_name  
## Kruskal-Wallis chi-squared = 2762.1, df = 3, p-value < 2.2e-16

kruskal.test(data = commute\_areas, trip\_speed ~ area\_name)

##   
## Kruskal-Wallis rank sum test  
##   
## data: trip\_speed by area\_name  
## Kruskal-Wallis chi-squared = 7659.1, df = 4, p-value < 2.2e-16

# run dunn test for determining rank differences of travel efficiencies across case study areas

NemenyiTest(x=walk\_areas$trip\_speed, g=walk\_areas$area\_name)

##   
## Nemenyi's test of multiple comparisons for independent samples (tukey)   
##   
## mean.rank.diff pval   
## Nordhavn-Mjølnerparken 321.6805 6.3e-05 \*\*\*  
## Taastrupgaard-Mjølnerparken -615.6737 4.0e-14 \*\*\*  
## Tingbjerg-Mjølnerparken 1055.0132 < 2e-16 \*\*\*  
## Taastrupgaard-Nordhavn -937.3543 < 2e-16 \*\*\*  
## Tingbjerg-Nordhavn 733.3327 4.7e-14 \*\*\*  
## Tingbjerg-Taastrupgaard 1670.6870 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

NemenyiTest(x=bike\_areas$trip\_speed, g=bike\_areas$area\_name)

##   
## Nemenyi's test of multiple comparisons for independent samples (tukey)   
##   
## mean.rank.diff pval   
## Nordhavn-Mjølnerparken 790.0047 4.0e-14 \*\*\*  
## Taastrupgaard-Mjølnerparken 425.8130 3.3e-08 \*\*\*  
## Tingbjerg-Mjølnerparken 2200.6692 < 2e-16 \*\*\*  
## Taastrupgaard-Nordhavn -364.1917 3.7e-06 \*\*\*  
## Tingbjerg-Nordhavn 1410.6645 < 2e-16 \*\*\*  
## Tingbjerg-Taastrupgaard 1774.8562 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

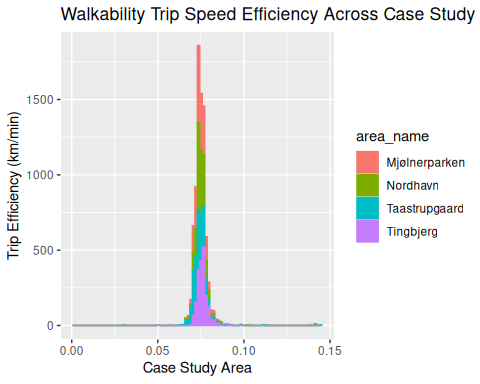
NemenyiTest(x=local\_transit\_areas$trip\_speed, g=local\_transit\_areas$area\_name)

##   
## Nemenyi's test of multiple comparisons for independent samples (tukey)   
##   
## mean.rank.diff pval   
## Nordhavn-Mjølnerparken -1706.7400 < 2e-16 \*\*\*  
## Taastrupgaard-Mjølnerparken -2673.7595 < 2e-16 \*\*\*  
## Tingbjerg-Mjølnerparken 761.1745 4.2e-14 \*\*\*  
## Taastrupgaard-Nordhavn -967.0195 < 2e-16 \*\*\*  
## Tingbjerg-Nordhavn 2467.9145 < 2e-16 \*\*\*  
## Tingbjerg-Taastrupgaard 3434.9340 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

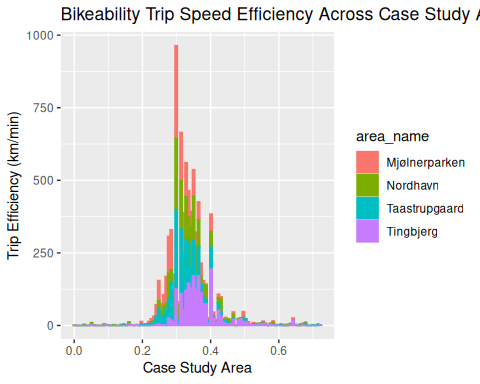
NemenyiTest(x=commute\_areas$trip\_speed, g=commute\_areas$area\_name)

##   
## Nemenyi's test of multiple comparisons for independent samples (tukey)   
##   
## mean.rank.diff pval   
## Nordhavn-Mjølnerparken -3360.8360 < 2e-16 \*\*\*  
## Taastrupgaard-Mjølnerparken 425.6070 3.1e-05 \*\*\*  
## Tingbjerg-Mjølnerparken -6494.8865 < 2e-16 \*\*\*  
## Tingbjerg by Train-Mjølnerparken -3138.4932 < 2e-16 \*\*\*  
## Taastrupgaard-Nordhavn 3786.4430 < 2e-16 \*\*\*  
## Tingbjerg-Nordhavn -3134.0505 < 2e-16 \*\*\*  
## Tingbjerg by Train-Nordhavn 222.3427 0.1059   
## Tingbjerg-Taastrupgaard -6920.4935 < 2e-16 \*\*\*  
## Tingbjerg by Train-Taastrupgaard -3564.1002 < 2e-16 \*\*\*  
## Tingbjerg by Train-Tingbjerg 3356.3933 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

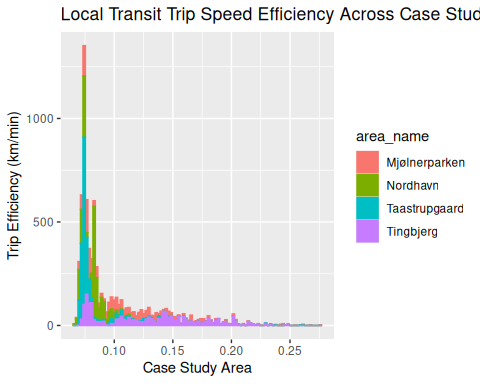
ggplot(data = walk\_areas, aes(x = trip\_speed, fill = area\_name, color = area\_name)) +   
 geom\_histogram(bins = 100) +   
 xlab("Case Study Area") +   
 ylab("Trip Efficiency (km/min)") +   
 labs(title = "Walkability Trip Speed Efficiency Across Case Study Areas")



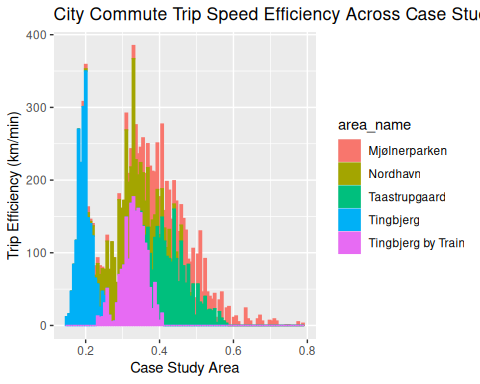
ggplot(data = bike\_areas, aes(x = trip\_speed, fill = area\_name, color = area\_name)) +   
 geom\_histogram(bins = 100) +   
 xlab("Case Study Area") +   
 ylab("Trip Efficiency (km/min)") +   
 labs(title = "Bikeability Trip Speed Efficiency Across Case Study Areas")



ggplot(data = local\_transit\_areas, aes(x = trip\_speed, fill = area\_name, color = area\_name)) +   
 geom\_histogram(bins = 100) +   
 xlab("Case Study Area") +   
 ylab("Trip Efficiency (km/min)") +   
 labs(title = "Local Transit Trip Speed Efficiency Across Case Study Areas")



ggplot(data = commute\_areas, aes(x = trip\_speed, fill = area\_name, color = area\_name)) +   
 geom\_histogram(bins = 100) +   
 xlab("Case Study Area") +   
 ylab("Trip Efficiency (km/min)") +   
 labs(title = "City Commute Trip Speed Efficiency Across Case Study Areas")



ggplot(data = alt\_commute\_areas, aes(x = trip\_speed, fill = area\_name, color = area\_name)) +   
 geom\_histogram(bins = 100) +   
 xlab("Case Study Area") +   
 ylab("Trip Efficiency (km/min)") +   
 labs(title = "Alt City Commute Trip Speed Efficiency Across Case Study Areas")

