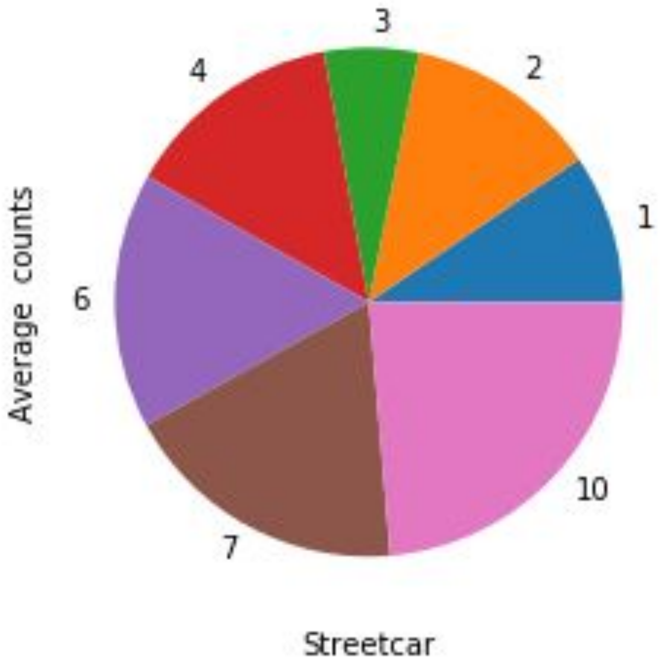
POISSON REGRESSION

EXPLORATIVE STATISTICS

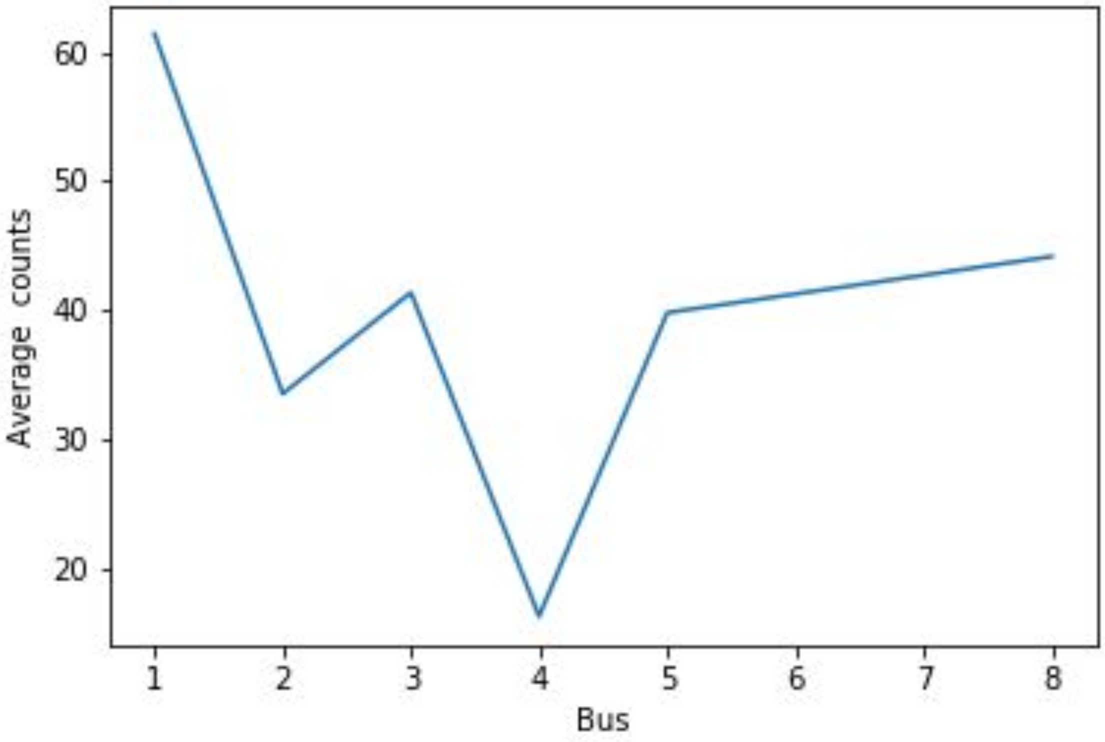
1. Relationships between infrastructures and checkouts
2. Distribution of checkouts

RELATIONSHIPS BETWEEN INFRASTRUCTURES AND CHECKOUTS

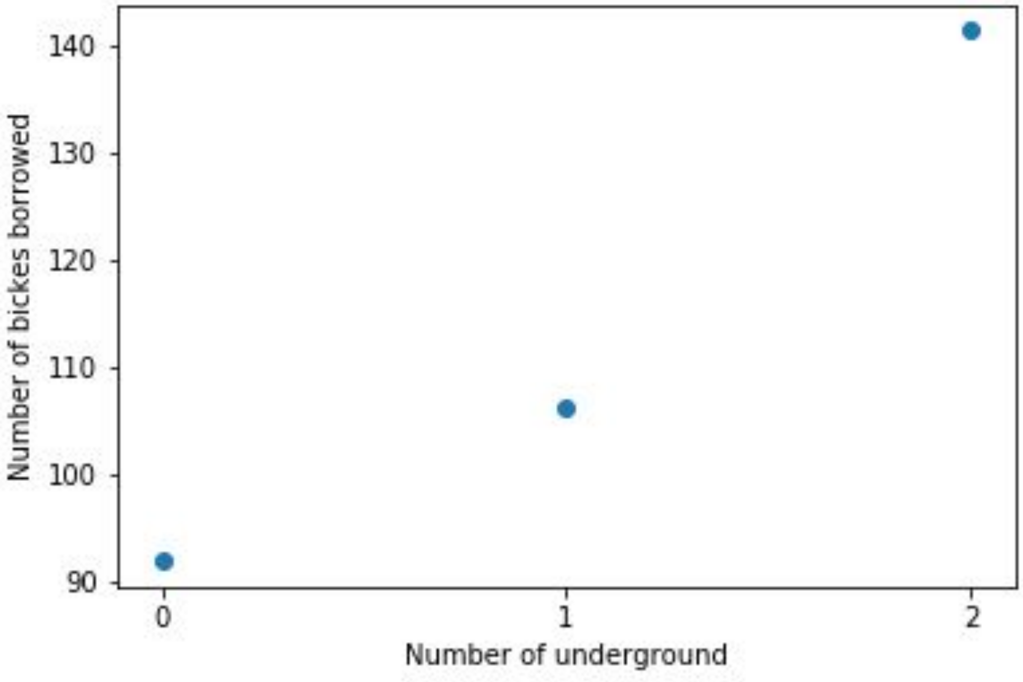
STREETCARS AND CHECKOUTS



BUSES AND CHECKOUTS



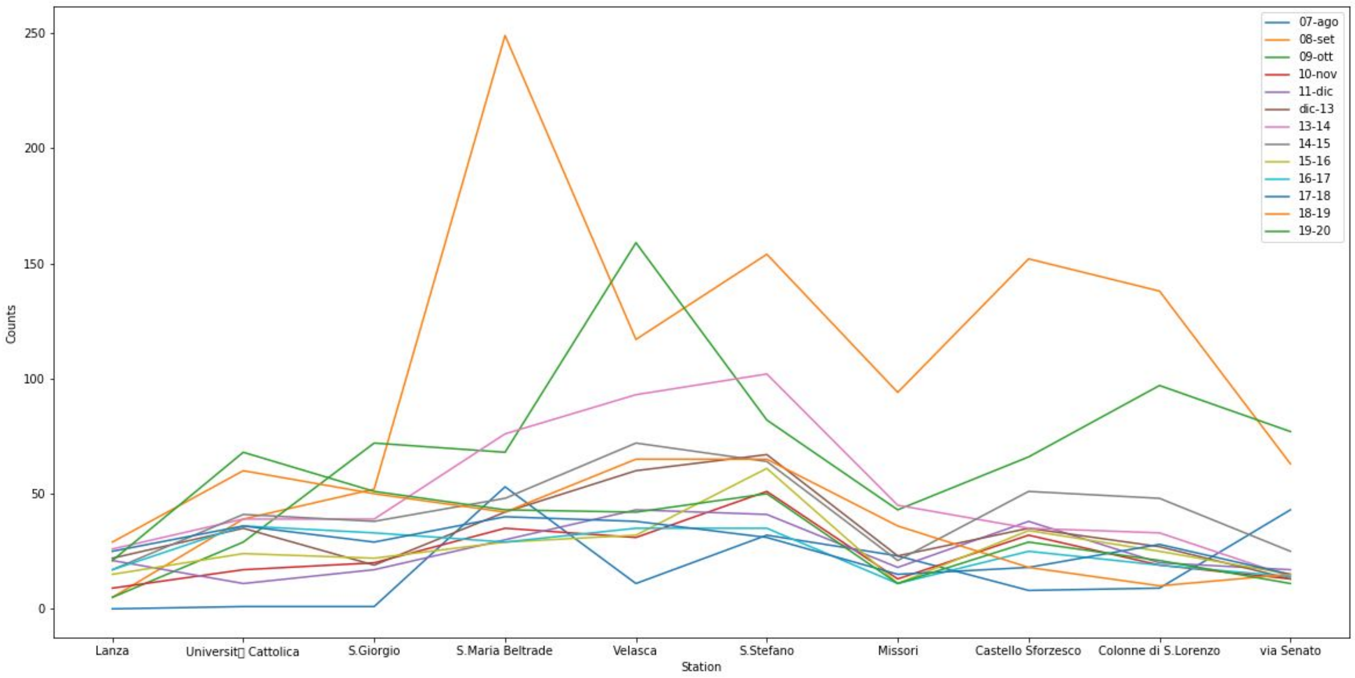
UNDERGROUND AND CHECKOUTS



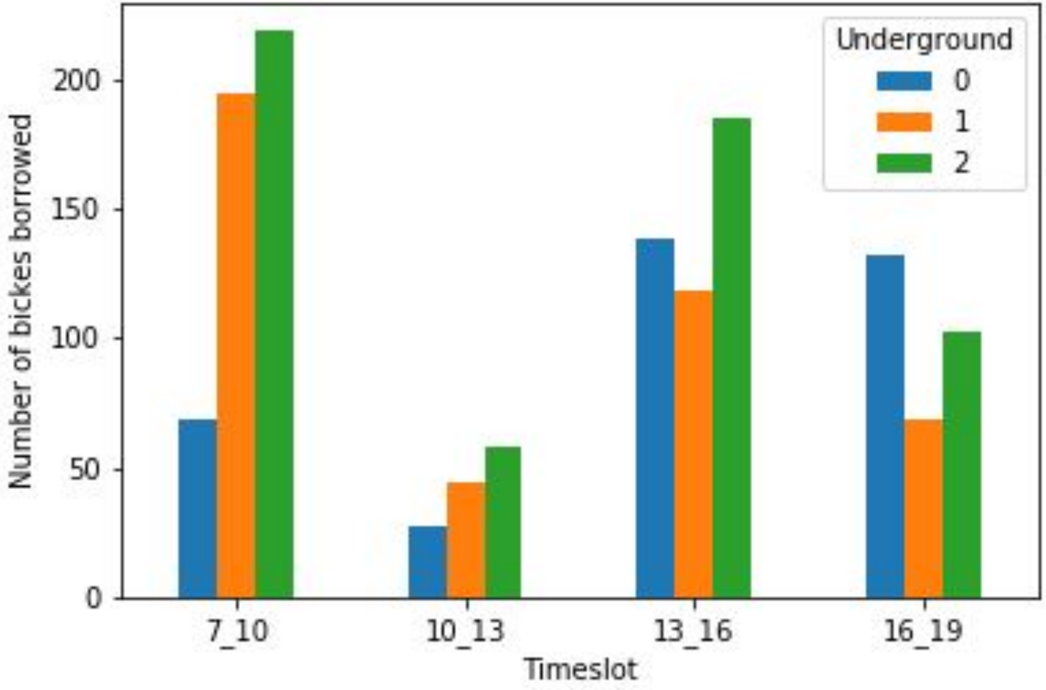
IDEAS

1. There is a clear and positive correlation between the number of undergrounds and the number of checkouts
2. Other infrastructures provide no interesting informations

CHECKOUTS IN EACH LOCATION



DISTRIBUTION OF CHECKOUTS



FREQUENTIST ANALYSIS

R RESULTS

```
Call:
glm(formula = y ~ x0 + x1 + x2, family = poisson(link = "log"),
    data = data)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-7.9530  -4.6717  -0.6827   4.0725   7.6495

Coefficients: (1 not defined because of singularities)
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  4.95053    0.04207 117.674  < 2e-16 ***
x0          -0.42874    0.06699  -6.400 1.55e-10 ***
x1          -0.28709    0.06425  -4.468 7.89e-06 ***
x2                NA         NA      NA       NA
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 394.01  on 11  degrees of freedom
Residual deviance: 349.40  on  9  degrees of freedom
AIC: 432.21

Number of Fisher Scoring iterations: 4
```

JAGS RESULTS

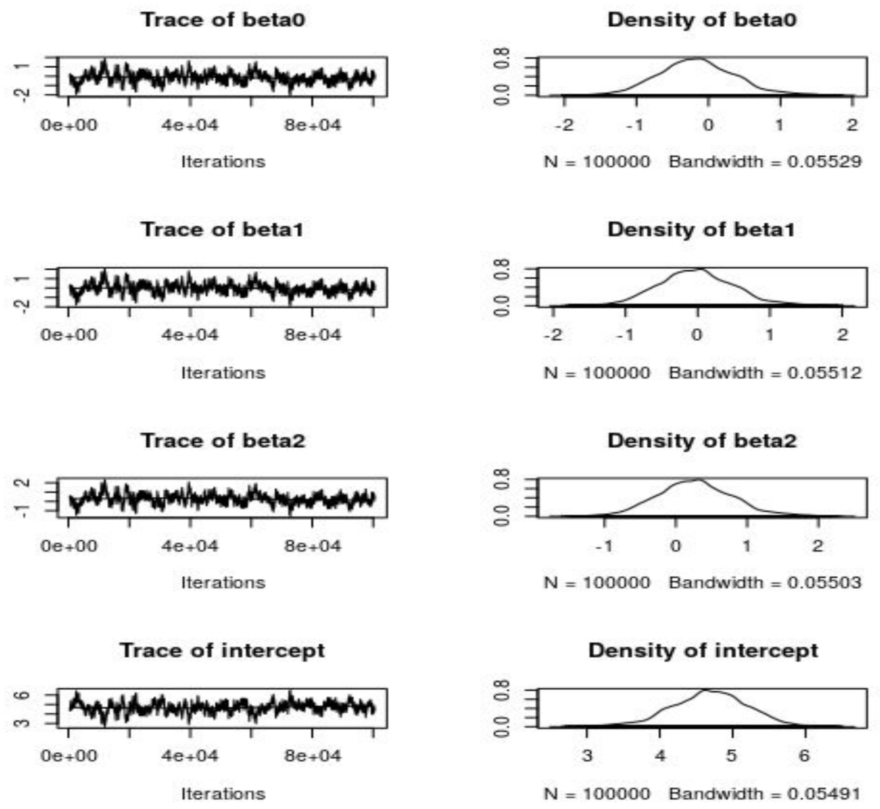
```
Iterations = 501:100500
Thinning interval = 1
Number of chains = 1
Sample size per chain = 1e+05

1. Empirical mean and standard deviation for each variable,
   plus standard error of the mean:

              Mean     SD Naive SE Time-series SE
beta0     -0.16048 0.5366 0.001697        0.05851
beta1     -0.01885 0.5361 0.001695        0.05837
beta2      0.26829 0.5364 0.001696        0.05868
intercept  4.68107 0.5356 0.001694        0.05870

2. Quantiles for each variable:

             2.5%      25%      50%    75%  97.5%
beta0     -1.1760 -0.51543 -0.17486 0.1835 0.9806
beta1     -1.0297 -0.37242 -0.03164 0.3244 1.1240
beta2     -0.7416 -0.08526  0.25638 0.6104 1.4084
intercept  3.5394  4.33978  4.69416 5.0339 5.6913
```



RESULTS

1. Very similar intercept
2. Slightly different slopes

BAYESIAN ANALYSIS

JAGS MODEL

```
model{
    for ( i in 1:12) {
      y[i] ~ dpois(mu[i])
      log(mu[i]) <- intercept + beta0*x0[i] + beta1*x1[i] + beta2*x2[i] + e
    }
    #priors
    intercept ~ dnorm(4,0.01)
    beta0 ~  dnorm(0,1)
    beta1 ~ dnorm(0,1)
    beta2 ~  dnorm(0,1)
    e ~ dnorm(0,1/lambda)
    #hyperpriors
    lambda ~ dgamma(3,2)

    }
```

PRIORS

1. Use normal distribution to model the approximation made by the optimization procedure and numerical approximation
2. Use gamma distribution as variance hyperprior

JAGS RESULTS

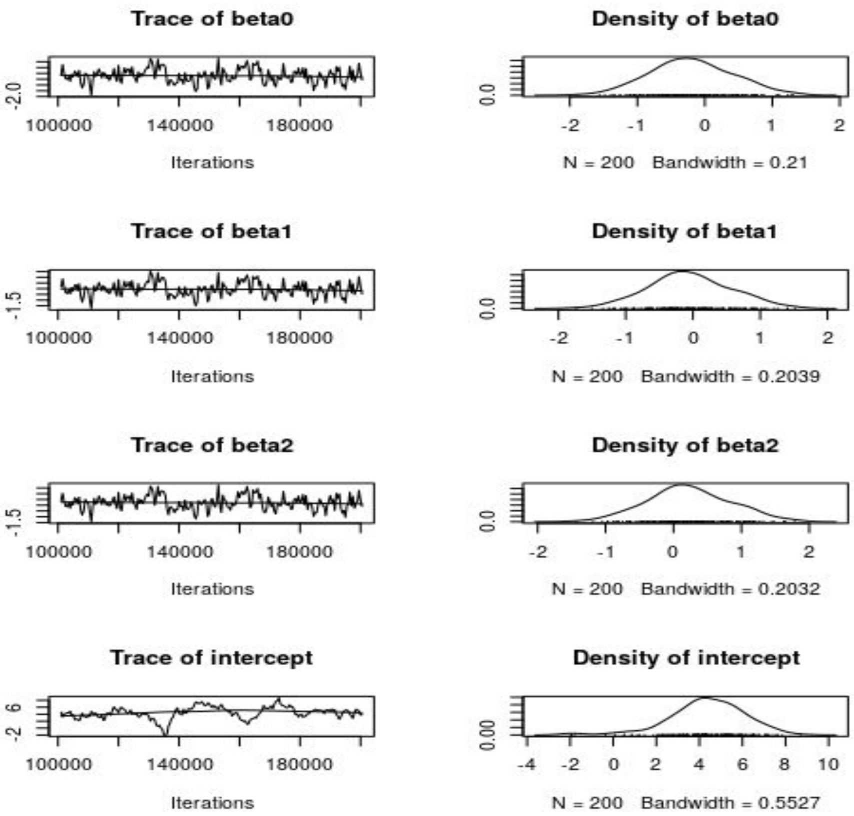
```
Iterations = 1000:100500
Thinning interval = 500
Number of chains = 1
Sample size per chain = 200

1. Empirical mean and standard deviation for each variable,
   plus standard error of the mean:

              Mean     SD Naive SE Time-series SE
beta0     -0.33213 0.6191  0.04378        0.07114
beta1     -0.19660 0.6243  0.04415        0.07232
beta2      0.09142 0.6187  0.04375        0.07225
e         -0.61515 1.1405  0.08064        0.22806
intercept  5.47270 1.3472  0.09526        0.29295

2. Quantiles for each variable:

            2.5%     25%     50%     75%  97.5%
beta0     -1.539 -0.7368 -0.3248 0.07618 0.8858
beta1     -1.374 -0.6422 -0.1658 0.25460 0.9955
beta2     -1.121 -0.3188  0.1254 0.53513 1.2550
e         -2.877 -1.4254 -0.5245 0.17661 1.2400
intercept  3.427  4.4427  5.3148 6.55345 8.0522
```



RESULTS

1. Very similar intercept w.r.t jags frequentist model
2. Slopes more similar to R regression model