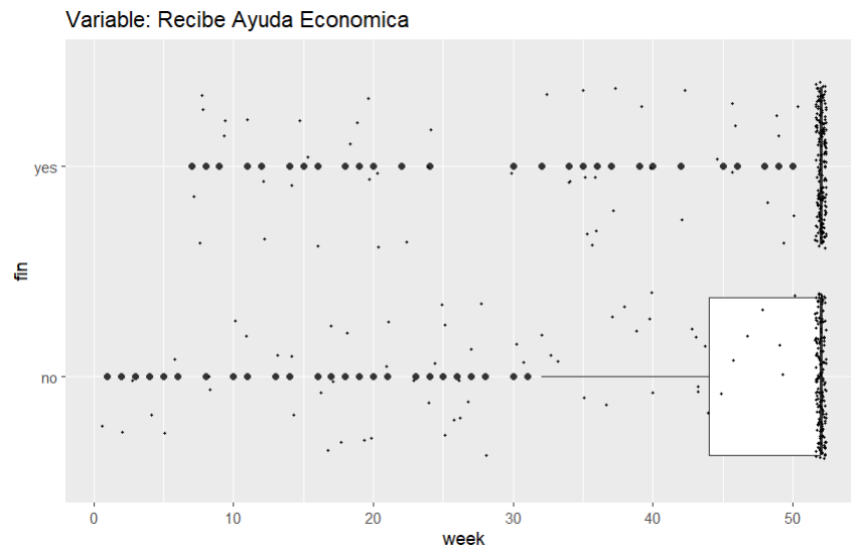


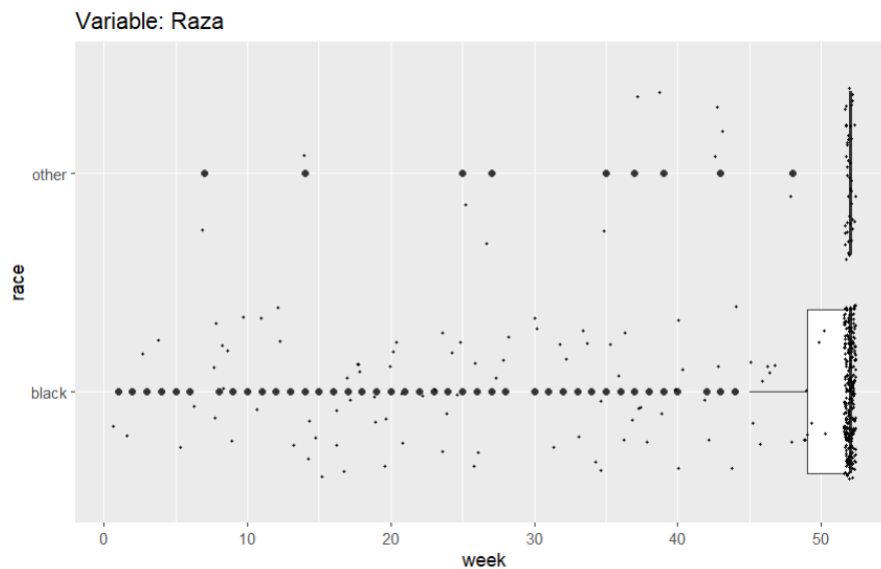
Respuestas tarea 1

Pregunta 1

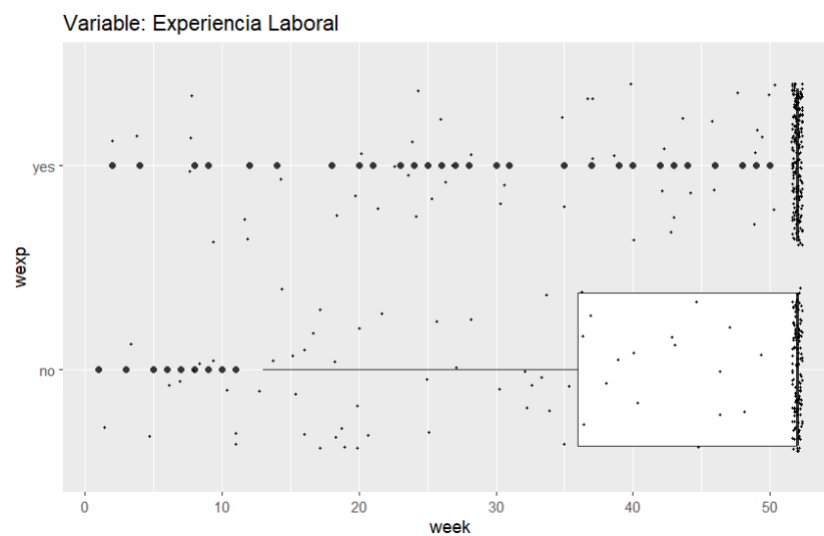
Analizamos las variables dadas en el dataset usando boxplots:



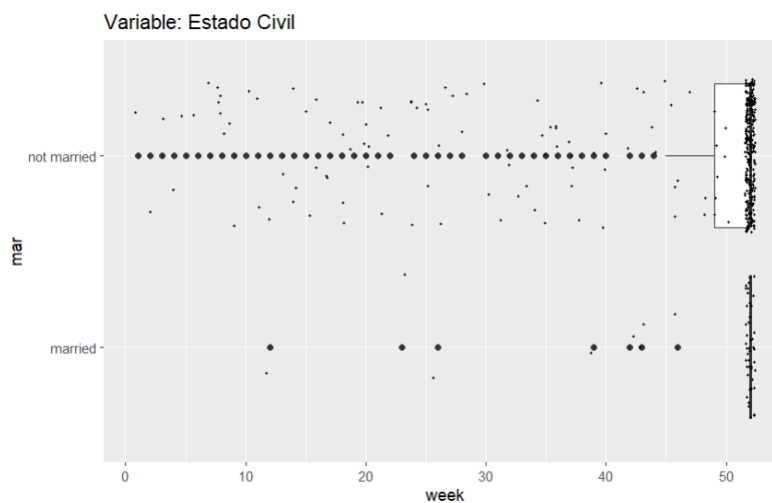
Este boxplot nos da la impresión de que las personas que no reciben ayuda económica tienden a ser arrestados en tiempos menores (más temprano).



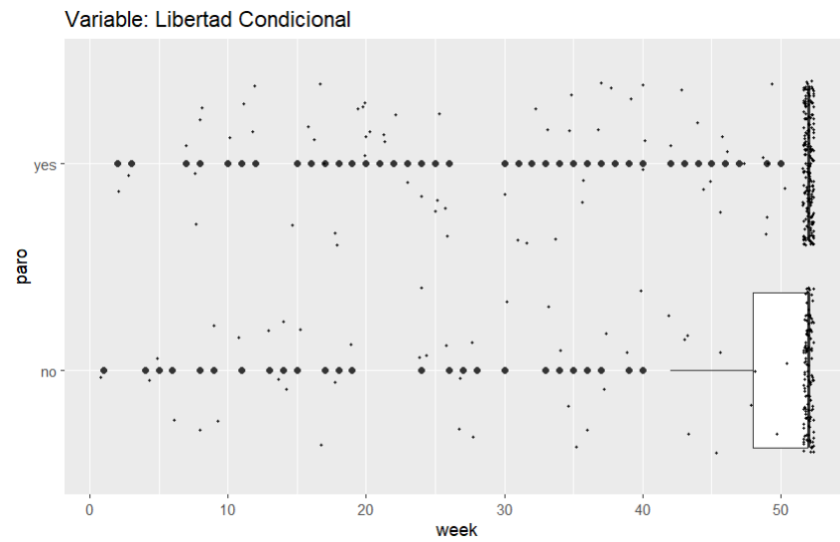
La gente de raza negra tienda a ser arrestados antes que otras razas. Sin embargo, hay que evaluar cómo se condujo la obtención de los datos.



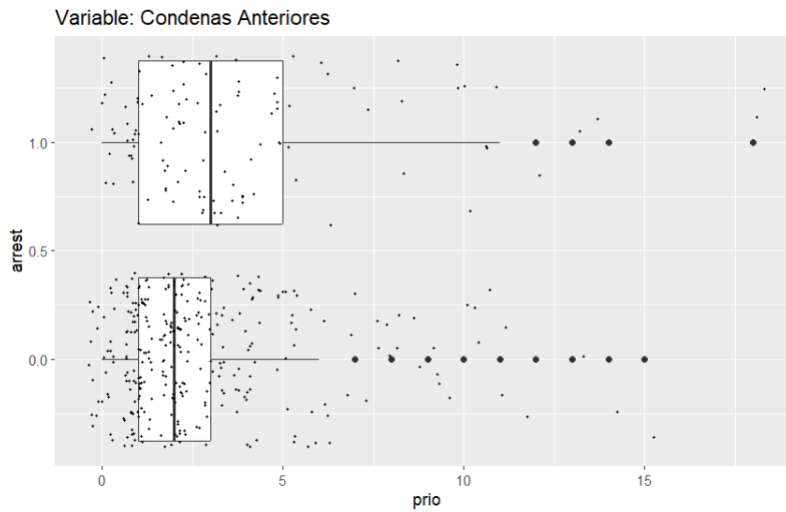
Las personas sin experiencia laboral tienen mayor probabilidad de ser arrestadas temprano en el periodo del experimento.



Los casados tienen menor probabilidad de ser arrestadas temprano.



Las personas con Libertad Condicional son menos propensas a ser arrestadas



Las personas que son arrestadas tienden a tener más condenas previas

Pregunta 2

A continuación, se presentan los Log Rank tests de cada variable individual:

Variable	p-value
Fin	0.05
Race	0.4
Wexp	0.002
Mar	0.05
Paro	0.6
Prio	0.0004
educ	0.03

Con esta información, las variables que generan curvas de sobrevivencia significativamente diferentes son: fin, wexp, mar, prio, educ.

Juntando las variables categóricas significativas, el Log Rank test da un valor-p de 0.04, lo que nos dice que esta combinación de variables genera curvas de sobrevivencia significativamente diferentes

Pregunta 3

Considerando todos los predictores, tenemos los siguientes resultados:

```

#> A logit model with survival as the
#> response variable.
#>
#>      coef exp(coef) se(coef)      z Pr(>|z|)
#>-----
fines      -0.40212   0.66890  0.19140 -2.101  0.0356 *
raceother  -0.29749   0.74268  0.30805 -0.966  0.3342
wexpyes    -0.31322   0.73109  0.20428 -1.533  0.1252
marnot married 0.54939   1.73219  0.37959  1.447  0.1478
paroyes    -0.02729   0.97308  0.19411 -0.141  0.8882
prio        0.08269   1.08621  0.02881  2.870  0.0041 **
educ       -0.17087   0.84293  0.12575 -1.359  0.1742
#>
#> ---
```

Las variables significativas son prio y fin.

Pregunta 4 y 5

Ajustando un modelo de Cox únicamente con las variables de prio y fin, podemos concluir que prio (el número de condenas anteriores) es un factor de riesgo pues tiene un valor exponencial de 1.11 y el fin (si la persona recibió ayuda financiera después de salir de prisión) es un factor protector pues tiene un valor exponencial de 0.67 (al ser menor de 0, reduce el riesgo).

```
      coef exp(coef) se(coef)      z Pr(>|z|)
finyes -0.39880   0.67112  0.19005 -2.098   0.0359 *
prio    0.10411   1.10972  0.02675  3.892 9.96e-05 ***
```

Pregunta 6

La prueba de hipótesis para validar el supuesto de riesgos proporcionales da un valor-p de 0.75, lo que nos indica que los riesgos sí son proporcionales:

```
> test
      chisq df    p
fin     0.0741  1 0.79
prio    0.5093  1 0.48
GLOBAL  0.5654  2 0.75
```