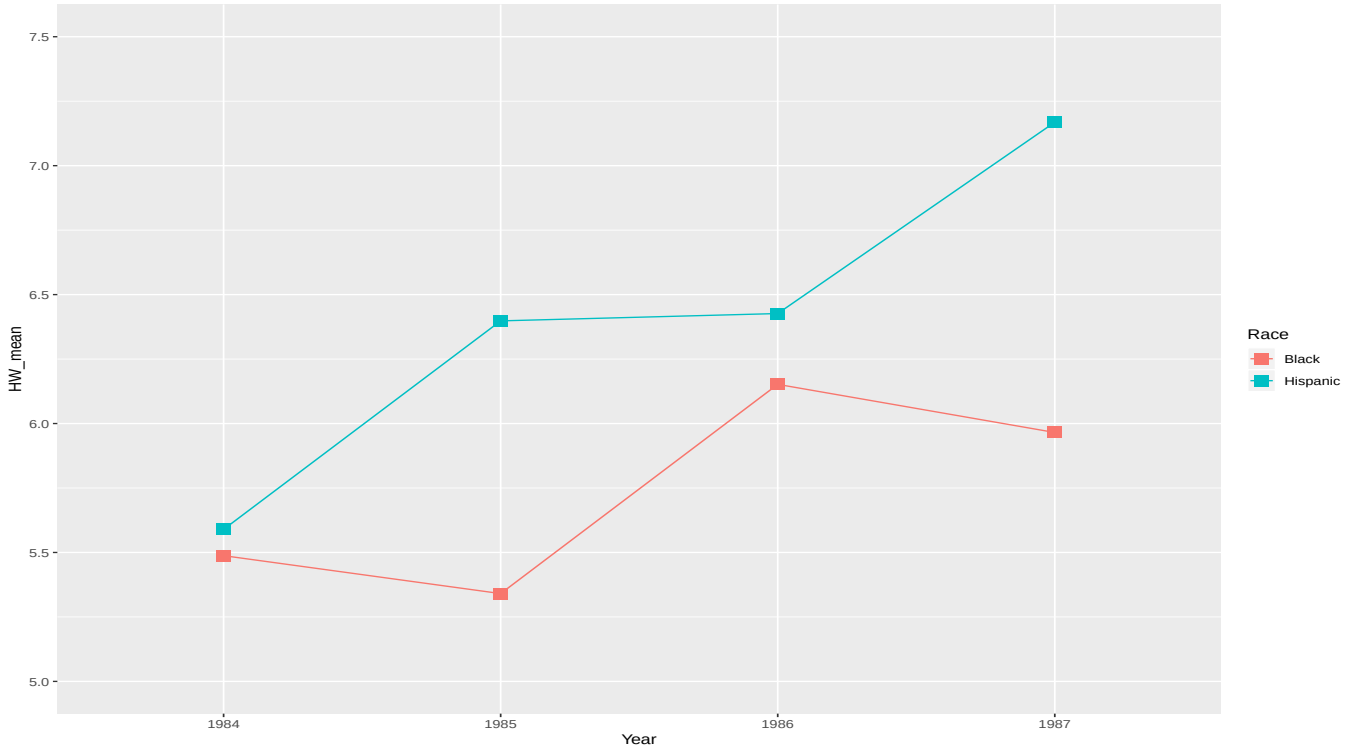
1 Graphical analysis

1. **(3 marks)**

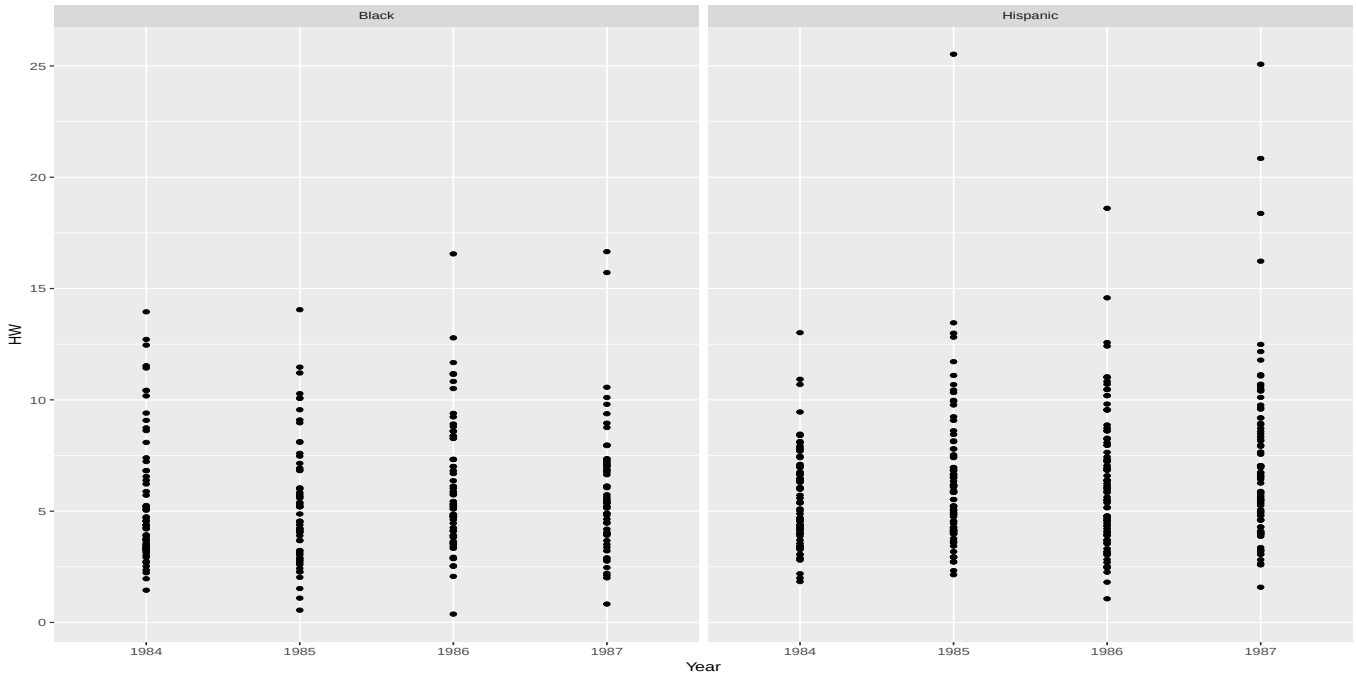
Figure 1: Plot of mean HW vs $Year$ grouped by $Race$



Yes since the trajectory of mean HW over time for the *Black* group is different to the *Hispanic* group. **(3 marks)**

2. **(3 marks)**

Figure 2: Scatter plot of HW vs $Year$ grouped by $Race$



It seems like the variability of HW is fairly constant across the years, with a possible exception of 1987 in the *Hispanic* group. **(3 marks)**. It seems that the variability of HW for each year doesn't seem to change across races, with a possible exception of 1987. **(3 marks)**

2 Describing the model

3. β_0 is the mean hourly wage for *Black* male workers in 1984. **(3 marks)**

4. • β_1 is the difference in mean hourly wage between 1985 and 1984, for *Black* male workers. **(3 marks)**

   • β_2 is the difference in mean hourly wage between 1986 and 1984, for *Black* male workers. **(3 marks)**

   • β_3 is the difference in mean hourly wage between 1987 and 1984, for *Black* male workers. **(3 marks)**

   • β_4 is the difference in mean hourly wage between *Hispanic* and *Black* male workers in 1984. **(3 marks)**

5. • β_5 is the change in the difference in mean hourly wage between *Hispanic* and *Black* male workers when we change from 1984 to 1985. or β_5 is the change in the difference in mean hourly wage between the years 1985 and 1984 when we change from *Black* to *Hispanic* male workers. **(3 marks)**

   • β_6 is the change in the difference in mean hourly wage between *Hispanic* and *Black* male workers when we change from 1984 to 1986. or β_6 is the change in the difference in mean hourly wage between the years 1986 and 1984 when we change from *Black* to *Hispanic* male workers. **(3 marks)**

   • β_7 is the change in the difference in mean hourly wage between *Hispanic* and *Black* male workers when we change from 1984 to 1987. or β_7 is the change in the difference in mean hourly wage between the years 1987 and 1984 when we change from *Black* to *Hispanic* male workers. **(3 marks)**

6. (a) **(2 marks)**

$$\boldsymbol{Y}_1 = \begin{bmatrix} 3.3738 \\ 3.2346 \\ 3.8207 \\ 3.9809 \end{bmatrix}.$$

   (b) **(4 marks)**

$$\boldsymbol{X}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

   (c) **(2 marks)**

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \\ \beta_7 \end{bmatrix}.$$

   (d) **(2 marks)**

$$\boldsymbol{\varepsilon}_i = \begin{bmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \\ \varepsilon_{3i} \\ \varepsilon_{4i} \end{bmatrix}.$$

3 Choosing the appropriate R matrix

7. (a)

Model (1A) AIC = 2625.803 **(1 mark)**
Model (1B) AIC = 2716.531 **(1 mark)**
Model (1C) AIC = 2657.499 **(1 mark)**

(b)

Model (1A) BIC = 2691.555 **(1 mark)**
Model (1B) BIC = 2760.366 **(1 mark)**
Model (1C) BIC = 2701.334 **(1 mark)**

(c) I would choose model (1A) since the AIC and BIC values for model (1A) are smaller than the AIC and BIC values of models (1B) and (1C). **(3 marks)**

4 Fixed effect estimates for your final linear mixed model

8. **(2 marks)**

Figure 3: Table 1

	Value	Std.Error	t-value	p-value
(Intercept)	5.4876	0.3798	14.4482	0.0000
Y2	-0.1471	0.3271	-0.4496	0.6532
Y3	0.6643	0.2873	2.3126	0.0211
Y4	0.4777	0.3557	1.3428	0.1798
H	0.1019	0.5012	0.2033	0.8390
Y2:H	0.9557	0.4316	2.2140	0.0272
Y3:H	0.1727	0.3791	0.4556	0.6489
Y4:H	1.1013	0.4694	2.3462	0.0193

9. Yes since $\hat{\beta}_5 = 0.9557$ which suggests that the difference in mean hourly wage between *Hispanic* and *Black* male workers increased by \$0.9557 from 1984 to 1985 and this increase was significant (p-value $= 0.0272 < 0.05$). **(5 marks)**

10. No since even though $\hat{\beta}_6 = 0.1727$ which suggests that the difference in mean hourly wage between *Hispanic* and *Black* male workers increased by \$0.1727 from 1984 to 1986, this increase was insignificant (p-value $= 0.6489 > 0.05$). **(5 marks)**

11. $\hat{\tau}_1 - \hat{\tau}_2 = -0.7830$ **(3 marks)**

12. Yes since $\hat{\tau}_1 - \hat{\tau}_2 = -0.7830$ which suggests that the difference in mean hourly wage between *Hispanic* and *Black* male workers decreased by \$0.7830 from 1985 to 1986 and this decrease was significant (p-value $= 0.0455 < 0.05$). **(5 marks)**

13. $\hat{\eta}_1 - \hat{\eta}_2 = 0.1456$ **(3 marks)**

14. No, since even though $\hat{\eta}_1 - \hat{\eta}_2 = 0.1456$ which suggests that the increase in mean hourly wage from 1985 to 1987 for *Hispanic* male workers was 0.1456 higher than the increase for *Black* male workers, this increase was insignificant (p-value $= 0.657 > 0.05$). **(5 marks)**