Introduction :
Capital Asset Pricing (also known as CAPM) is a model that helps to show the relationship between risk and the expected return for assets (in this particular case; stocks). Essentially; it is a formula created through the use of 3 key ‘variables’:

- Risk free rate : represents the time-value of money and is typically the yield on government bonds
- Beta of the security : a measure of systematic risk of a portfolio in comparison to the market
- Market Risk Premium

In terms of when it is used; it is used in a wide variety of possible investing situations including :

- The pricing of risky securities
- Calculating costs of capital
- Expected returns for assets; given the risk of those particular assets

Advantages of CAPM:

- It considers only systematic risk given that investors have diversified their portfolios
- It is more reliable and effective method of calculating risk than other models.
- CAPM takes into account a company’s level of systematic risk against the stock market as a whole. This is beneficial as it allows for a company to compare itself to the market as long as we have Beta

Disadvantages of CAPM:

- CAPM is largely based on assumptions such as estimating market risk premium and beta, which can vary over time
- We are depending on the past to predict future for Beta and that again isn’t always reliable

In comparison, the Fama and French Model is a 3 factor model that expands on the CAPM. It is said to be a better tool for evaluating manager performance due to adding size and and value factors to the CAPM model. Thus it allows for ‘outperformance tendency’ which considers for the idea that small market capitalisation stocks outperform markets on a regular basis, which helps to predict better return and thus better results.

The goal of this regression is to determine which types of risks are vital in order to determine the price of an asset (or in this case, portfolio). Further, the regression attempts to help provide a discussion as to whether or not price can be affected by something other than risk, such as seasonal trends. Further, it attempts to discuss whether a model with more variables provides more cohesive information.

Data :

The data used for this report has been collected from Kenneth R. Frenchís data library, at the following website :
http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
It contains information about a range of industries and their relative portfolio's using information gathered through the CAPM and Fama and French 3 factor model. This information has been compiled into spreadsheets based on years and models in order to allow for analysis and discussion. An industry portfolio is essentially a grouping of assets (in this case stocks) and should be constructed with information relating to their risk tolerance and investing objectives.
The 5 industry portfolios that are 'discussed' are:

- **Consumer** : relate to the items purchased by individuals rather than manufacturers or companies.
- **Manufacturing** : relates to the stocks of companies which manufacture (produce) a product on a large scale using machinery
- **High Tech** : relate to the stocks of companies operating in high technology fields which use advanced methods and the most modern equipment
- **Health Care** : relates to stocks of companies that provide medical services and manufacture medical equipment or drugs and the like.
- **Other** : Includes industries which do not fall under a specific industry portfolio such as sanitary services.

These are from the stocks :

- **NYSE** : New York Stock Exchange
- **AMEX** : American Stock Exchange
- **NASDAQ** : National Association of Securities Dealers Automated Quotations

PART A

Interpretation : We want the model with the smallest BIC, HQ or AIC and the largest adjusted R squared

CAPM

$\widehat{E(R)} = R_f + \beta(R_m - R_f)$

where $\widehat{E(R)}$ is the estimated market return, R_f is the risk free rate and $R_m - R_f$ is the market risk premium

FAMA FRENCH 3 FACTOR MODEL

$\widehat{E(R)} = R_f + \beta_1(R_m - R_f) + \beta_2 SMB + \beta_3 HML$

where $\widehat{E(R)}$ is the estimated market return, R_f is the risk free rate, $R_m - R_f$ is the market risk premium, SMB (Small Minus Big) is the average return on the three small portfolios minus the average return on the three big portfolios and HML (High Minus Low) is the average return on the two value portfolios minus the average return on the two growth portfolios

We will be using Fama French and 5 industry portfolio csv files to conduct our analysis. We will be using monthly data from the Fama french csv and in the 5 industry portfolio csv, we picked the returns for each industry that are equally weighted. We merged these two csv files based on their date.

Industry 1: Consumer

CAPM

$$\widehat{E(R)} - R_f = \underset{(0.104)}{0.124} + \underset{(0.0194)}{1.17}\,(R_m - R_f)$$

Interpretation of each estimated parameter:

$\boldsymbol{\beta_0}$: When expected return on the market rate is equal to the risk free rate ($R_m = R_f$), the additional return for market risk will be 0.124%

$\boldsymbol{\beta_1}$: An increase in market premium by 1% will result in a 1.17% increase in the additional return for market risk, keeping all else constant.

> This does not make sense because the beta of the security ($\boldsymbol{\beta_1}$) is greater than 1 (i.e. 1.17%) hence it is deemed to be risky. However, consumer industry will be quite stable and hence it doesn't make sense for it to be risky. As a result, the additional return for market risk will not increase by more than 1% as $\boldsymbol{\beta_1}$ should be less than 1%.

Hypothesis Testing:

$H_0 : \beta_1 = 0$
$H_1 : \beta_1 \neq 0$

β_1 : Since, $|t_{calc}| > t_{crit}$ (60.26 > 1.960), we reject the null hypothesis (H_0) at 5% significance level. Therefore, market risk premium holds statistical significance at 5% level of significance.

Assessing the residuals:

We need to check if there are Heteroskedastic error terms for this model. We will be using Breusch pagan test to test for Heteroskedasticity.

H_0 : $E(u_i^2) = \sigma^2$ (homoskedasticity)
H_1 : $E(u_i^2) \neq \sigma^2$ (heteroskedasticity)

We reject the null because the value of the n $\times$ $R_{\hat{u}}^2$ here is 63.91 which is larger than 3.84, the 5% critical value of the χ_1^2 distribution. This proves that there is heteroskedasticity and we will be using HAC standard errors to get a new equation.

Thus, the new estimated equation based on HAC standard errors is,

$$\widehat{E(R)} - R_f = \underset{(0.116)}{0.124} + \underset{(0.0445)}{1.17}\,(R_m - R_f)$$

Even though, the estimated coefficients are the same, but the standard errors and t-statistics have changed.

We strongly reject the null hypothesis that Market risk premium is insignificant when using a t statistic based on either OLS standard errors or on HAC standard errors.

FAMA FRENCH

$$\widehat{E(R)} - R_f = \underset{(0.0602)}{-0.0289} + \underset{(0.0120)}{0.960}(R_m - R_f) + \underset{(0.0196)}{0.843}\,SMB + \underset{(0.0176)}{0.299}\,HML$$

<u>Interpretation of each estimated parameter:</u>

β_0: When expected return on the market rate is equal to the risk free rate ($R_m = R_f$), the additional return for market risk will be -0.0289%

β_1: A 1% increase in the market risk premium will result in a 0.960% increase in the additional return for market risk, keeping all else constant.

> This is plausible because the beta of the security (β_1) is less than 1 (i.e. 0.96%) hence it is not deemed to be risky. Since consumer industry will be quite stable and hence it makes sense that it has a low risk. As a result, the additional return for market risk will not increase by more than 1%.

β_2: A 1% increase in SMB will result in a 0.843% increase in the additional return for market risk, keeping all else constant.

> This is plausible because a positive SMB means that the portfolio has more small cap companies in it and if a portfolio has more small-cap companies in it, it should outperform the market over the long run therefore the additional return for market risk will increase.

β_3: A 1% increase in HML will result in a 0.299% increase in the estimated return, keeping all else constant.

> This is plausible because a positive HML means that the portfolio contains a large proportion of value stocks (High book to market ratio), a portfolio with a large proportion of value stocks should outperform one with a large proportion of growth

stocks (Low book to market ratio). Therefore, the additional return for market risk will increase.

Hypothesis Testing:

$H_0 : \beta_1 = 0$
$H_1 : \beta_1 \neq 0$
β_1 : Since, $|t_{calc}| > t_{crit}$ (79.81 > 1.960), we reject the null hypothesis (H_0) at 5% significance level. Therefore, market risk premium holds statistical significance at 5% level of significance.

$H_0 : \beta_2 = 0$
$H_1 : \beta_2 \neq 0$
β_2 : Since, $|t_{calc}| > t_{crit}$ (42.92 > 1.960), reject the null hypothesis (H_0) at 5% significance level. Therefore, SMB holds statistical significance at 5% level of significance.

$H_0 : \beta_3 = 0$
$H_1 : \beta_3 \neq 0$
β_3 : Since, $|t_{calc}| > t_{crit}$ (17.00 > 1.960), reject the null hypothesis (H_0) at 5% significance level. Therefore, HML holds statistical significance at 5% level of significance.

Assessing the residuals:

We need to check if there are Heteroskedastic error terms for this model. We will be using Breusch pagan test to test for Heteroskedasticity.

$H_0 : E(u_i^2) = \sigma^2$ (homoskedasticity)
$H_1 : E(u_i^2) \neq \sigma^2$ (heteroskedasticity)

We reject the null because the value of the n $\times$ $R_{\hat{u}}^2$ here is 56.57 which is larger than 7.81, the 5% critical value of the χ_3^2 distribution. This proves that there is heteroskedasticity and we will be using HAC standard errors to get a new equation.

Thus, the new estimated equation by using HAC standard errors is:

$$\widehat{E(R)} - R_f = -0.0289 + 0.960(R_m - R_f) + 0.843\,SMB + 0.299\,HML$$
$$(0.0695) \quad (0.0163) \quad (0.0750) \quad (0.0480)$$

Even though, the estimated coefficients are the same, but the standard errors and t-statistics have changed.

We strongly reject the null hypothesis that Market risk premium, SMB and HML are insignificant when using a t statistic based on either OLS standard errors or on HAC standard errors.

Industry 2: Manufacturing

CAPM

$\widehat{E(R)} - R_f = 0.160 + 1.23(R_m - R_f)$

(0.0949) (0.0176)

Interpretation of each estimated parameter:

$\boldsymbol{\beta_0}$**:** When expected return on the market rate is equal to the risk free rate ($R_m = R_f$), the additional return for market risk will be 0.160%

$\boldsymbol{\beta_1}$: A 1% increase in market risk premium will result in a 1.23% increase in the additional return for market risk, keeping all else constant.

> This is plausible because manufacturing industry usually has higher risk therefore the beta value ($\boldsymbol{\beta_1}$)will indeed be greater than 1 (i.e. 1.23). Hence, the additional return for market risk will increase by more than 1%.

Hypothesis Testing:

$H_0 : \beta_1 = 0$
$H_1 : \beta_1 \neq 0$

β_1 : Since, $|t_{calc}| > t_{crit}$ (69.66 > 1.960), we reject the null hypothesis (H_0) at 5% significance level. Therefore, Market risk premium holds statistical significance at 5% level of significance.

Assessing the residuals:

We need to check if there are Heteroskedastic error terms for this model. We will be using Breusch pagan test to test for Heteroskedasticity.

$H_0 : E(u_i^2) = \sigma^2$ (homoskedasticity)
$H_1 : E(u_i^2) \neq \sigma^2$ (heteroskedasticity)

We reject the null because the value of the n $\times$ $R_{\hat{u}}^2$ here is 54.90 which is larger than 3.84, the 5% critical value of the χ_1^2 distribution. This proves that there is heteroskedasticity and we will be using HAC standard errors to get a new equation.

Thus, the new estimated equation by using HAC standard errors is

$$\widehat{E(R)} - R_f = 0.160 + 1.23(R_m - R_f)$$
$$(0.0994) \quad (0.0461)$$

Even though, the estimated coefficients are the same, but the standard errors and t-statistics have changed.

We strongly reject the null hypothesis that Market risk premium are insignificant when using a t statistic based on either OLS standard errors or on HAC standard errors.

FAMA FRENCH

$$\widehat{E(R)} - R_f = -0.0159 + 1.04(R_m - R_f) + 0.663\,SMB + 0.435\,HML$$
$$(0.0548) \quad (0.0109) \qquad (0.0179) \qquad (0.0160)$$

<u>Interpretation of each estimated parameter:</u>

β_0: When expected return on the market rate is equal to the risk free rate ($R_m = R_f$), the additional return for market risk will be -0.0159%.

β_1: A 1% increase in the market risk premium will result in a 1.04% increase in the additional risk for market risk, keeping all else constant.

> This is plausible because manufacturing industry usually has higher risk therefore the beta value (β_1) will indeed be greater than 1 (i.e. 1.04). Hence, the additional return for market risk will increase by more than 1%.

β_2: A 1% increase in SMB will result in a 0.663% increase in the estimated return, keeping all else constant.

> This is plausible because a positive SMB means that the portfolio has more small cap companies in it and if a portfolio has more small-cap companies in it, it should outperform the market over the long run therefore the additional return for market risk will increase.

β_3: A 1% increase in HML will result in a 0.435% increase in the estimated return, keeping all else constant.

> This is plausible because a positive HML means that the portfolio contains a large proportion of value stocks (High book to market ratio), a portfolio with a large proportion of value stocks should outperform one with a large proportion of growth stocks (Low book to market ratio). Therefore, the additional return for market risk will increase.

<u>Hypothesis Testing:</u>

$H_0 : \beta_1 = 0$
$H_1 : \beta_1 \neq 0$
β_1 : Since, $|t_{calc}| > t_{crit}$ (94.66 > 1.960), we reject the null hypothesis (H_0) at 5% significance level. Therefore, Market risk premium holds statistical significance at 5% level of significance.

$H_0 : \beta_2 = 0$
$H_1 : \beta_2 \neq 0$
β_2 : Since, $|t_{calc}| > t_{crit}$ (37.08 > 1.960), we reject the null hypothesis (H_0) at 5% significance level. Therefore, SMB holds statistical significance at 5% level of significance.

$H_0 : \beta_3 = 0$
$H_1 : \beta_3 \neq 0$
β_3 : Since, $|t_{calc}| > t_{crit}$ (27.19 > 1.960), we do not reject the null hypothesis (H_0) at 5% significance level. Therefore, HML holds statistical significance at 5% level of significance.

Assessing the residuals:

We need to check if there are Heteroskedastic error terms for this model. We will be using Breusch pagan test to test for Heteroskedasticity.

$H_0 : E(u_i^2) = \sigma^2$ (homoskedasticity)
$H_1 : E(u_i^2) \neq \sigma^2$ (heteroskedasticity)

We reject the null because the value of the n $\times$ $R_{\hat{u}}^2$ here is 24.01 which is larger than 7.81, the 5% critical value of the χ_3^2 distribution. This proves that there is heteroskedasticity and we will be using HAC standard errors to get a new equation.

Thus, the new estimated equation by using HAC standard errors is:

$$\widehat{E(R)} - R_f = -0.0159 + 1.04(R_m - R_f) + 0.663\,SMB + 0.435\,HML$$
$$(0.0655) \quad (0.0172) \qquad (0.0529) \qquad (0.0377)$$

Even though, the estimated coefficients are the same, but the standard errors and t-statistics have changed.

We strongly reject the null hypothesis that Market risk premium, SMB and HML are insignificant when using a t statistic based on either OLS standard errors or on HAC standard errors.

Industry 3: High Tech

CAPM

$$\widehat{E(R)} - R_f = \underset{(0.128)}{0.228} + \underset{(0.0238)}{1.36}(R_m - R_f)$$

Interpretation of each estimated parameter:

β_0: When expected return on the market rate is equal to the risk free rate ($R_m = R_f$), the additional return for market risk will be 0.228%.

β_1: A 1% increase in market risk premium will result in a 1.36% increase in the additional return for market risk, keeping all else constant.

> This is plausible because High tech industry usually has higher risk therefore the beta value (β_1) will indeed be greater than 1 (i.e. 1.36). Hence, the additional return for market risk will increase by more than 1%.

Hypothesis Testing:

$H_0 : \beta_1 = 0$
$H_1 : \beta_1 \neq 0$

β_1 : Since, $|t_{calc}| > t_{crit}$ (57.23 > 1.960), we reject the null hypothesis (H_0) at 5% significance level. Therefore,Market risk premium holds statistical significance at 5% level of significance.

Assessing the residuals:

We need to check if there are Heteroskedastic error terms for this model. We will be using Breusch pagan test to test for Heteroskedasticity.

$H_0 : E(u_i^2) = \sigma^2$ (homoskedasticity)
$H_1 : E(u_i^2) \neq \sigma^2$ (heteroskedasticity)

We reject the null because the value of the n $\times$ $R_{\hat{u}}^2$ here is 4.90 which is larger than 3.84, the 5% critical value of the χ_1^2 distribution. This proves that there is heteroskedasticity and we will be using HAC standard errors to get a new equation.

Thus, the new estimated equation by using HAC standard errors is

$$\widehat{E(R)} - R_f = \underset{(0.140)}{0.228} + \underset{(0.0293)}{1.36}(R_m - R_f)$$

Even though, the estimated coefficients are the same, but the standard errors and t-statistics have changed.

We strongly reject the null hypothesis that Market risk premium are insignificant when using a t statistic based on either OLS standard errors or on HAC standard errors.

FAMA FRENCH

$$\widehat{E(R)} - R_f = \underset{(0.0931)}{0.201} + \underset{(0.0186)}{1.21}(R_m - R_f) + \underset{(0.0304)}{0.946}\, SMB - \underset{(0.0272)}{0.189}\, HML$$

Interpretation of each estimated parameter:

β_0: When expected return on the market rate is equal to the risk free rate ($R_m = R_f$), the additional return for market risk will be 0.201%.

β_1: A 1% increase in market risk risk premium will result in a 1.21% increase in the additional return for market risk, keeping all else constant.

> This is plausible because High tech industry usually has higher risk therefore the beta value (β_1) will indeed be greater than 1 (i.e. 1.21). Hence, the additional return for market risk will increase by more than 1%.

β_2: A 1% increase in SMB will result in a 0.946% increase in the estimated return, keeping all else constant.

> This is plausible because a positive SMB means that the portfolio has more small cap companies in it and if a portfolio has more small-cap companies in it, it should outperform the market over the long run therefore the additional return for market risk will increase.

β_3: A 1% increase in HML will result in a 0.189% decrease in the estimated return, keeping all else constant.

> This is plausible because a negative HML means that the portfolio contains a large proportion of growth stocks (Low book to market ratio), a portfolio with a large proportion of growth stocks will underperform as compared to the one with a large proportion of value stocks (High book to market ratio). Therefore, the additional return for market risk will decrease.

Hypothesis Testing:

$H_0 : \beta_1 = 0$
$H_1 : \beta_1 \neq 0$
β_1 : Since, $|t_{calc}| > t_{crit}$ (65.19 > 1.960), we reject the null hypothesis (H_0) at 5% significance level. Therefore, Market risk premium holds statistical significance at 5% level of significance.

$H_0 : \beta_2 = 0$
$H_1 : \beta_2 \neq 0$
β_2 : Since, $|t_{calc}| > t_{crit}$ (31.14 > 1.960), reject the null hypothesis (H_0) at 5% significance level. Therefore, SMB does not hold statistical significance at 5% level of significance.

$H_0 : \beta_3 = 0$
$H_1 : \beta_3 \neq 0$
β_3 : Since, $|t_{calc}| > t_{crit}$ (6.94 > 1.960), we reject the null hypothesis (H_0) at 5% significance level. Therefore, HML holds statistical significance at 5% level of significance.

<u>Assessing the residuals:</u>

We need to check if there are Heteroskedastic error terms for this model. We will be using Breusch pagan test to test for Heteroskedasticity.

$H_0 : E(u_i^2) = \sigma^2$ (homoskedasticity)
$H_1 : E(u_i^2) \neq \sigma^2$ (heteroskedasticity)

We reject the null because the value of the n $\times$ $R_{\hat{u}}^2$ here is 33.41 which is larger than 7.81, the 5% critical value of the χ_3^2 distribution. This proves that there is heteroskedasticity and we will be using HAC standard errors to get a new equation.

Thus, the new estimated equation by using HAC standard errors is

$$\widehat{E(R)} - R_f = \underset{(0.102)}{0.201} + \underset{(0.0315)}{1.21}(R_m - R_f) + \underset{(0.0764)}{0.946}\, SMB - \underset{(0.0724)}{0.189}\, HML$$

Even though, the estimated coefficients are the same, but the standard errors, t- statistics and p-values have changed.

We strongly reject the null hypothesis that Market risk premium,SMB and HML are insignificant when using a t statistic based on either OLS standard errors or on HAC standard errors.

<u>Industry 4: Healthcare</u>

CAPM

$$\widehat{E(R)} - R_f = 0.387 + 1.05(R_m - R_f)$$
$$(0.128) \quad (0.0238)$$

Interpretation of each estimated parameter:

β_0: When expected return on the market rate is equal to the risk free rate ($R_m = R_f$), the additional return for market risk will be 0.387%.

β_1: A 1% increase in market risk premium will result in a 1.05% increase in the additional return for market risk, keeping all else constant.

> This does not make sense because the beta of the security (β_1) is greater than 1 (i.e. 1.05%) hence it is deemed to be risky. However, healthcare industry will be quite stable and hence it doesn't make sense for it to be risky. As a result, the additional return for market risk will not increase by more than 1% as β_1 should be less than 1%.

Hypothesis Testing:

$H_0 : \beta_1 = 0$
$H_1 : \beta_1 \neq 0$

β_1 : Since, $|t_{calc}| > t_{crit}$ (44.39 > 1.960), we reject the null hypothesis (H_0) at 5% significance level. Therefore, Market risk premium holds statistical significance at 5% level of significance.

Assessing the residuals:

We need to check if there are Heteroskedastic error terms for this model. We will be using Breusch pagan test to test for Heteroskedasticity.

$H_0 :\ E(u_i^2) = \sigma^2$ (homoskedasticity)
$H_1 :\ E(u_i^2) \neq \sigma^2$ (heteroskedasticity)

We reject the null because the value of the n $\times$ $R_{\hat{u}}^2$ here is 4.15 which is larger than 3.84, the 5% critical value of the $\chi_1{}^2$ distribution. This proves that there is heteroskedasticity and we will be using HAC standard errors to get a new equation.

Thus, the new estimated equation by using HAC standard errors is

$\widehat{E(R)} - R_f =$ 0.387 + $1.05(R_m - R_f)$
(0.149) (0.0340)

Even though, the estimated coefficients are the same, but the standard errors and t-statistics have changed.

We strongly reject the null hypothesis that Market risk premium are insignificant when using a t statistic based on either OLS standard errors or on HAC standard errors

FAMA FRENCH

$\widehat{E(R)} - R_f =$ 0.392 + $0.945(R_m - R_f)$ + $0.785\, SMB$ - $0.257\, HML$
(0.103) (0.0207) (0.0337) (0.0302)

Interpretation of each estimated parameter:

β_0: When expected return on the market rate is equal to the risk free rate ($R_m = R_f$), the additional return for market risk will be 0.392%.

β_1: A 1% increase in market risk premium will result in a 0.945% increase in the additional return for market risk, keeping all else constant.

> This is plausible because the beta of the security (β_1) is less than 1 (i.e. 0.945%) hence it is deemed to be less risky. Since, healthcare industry will be quite stable and hence it makes sense for it to be less risky. As a result, the additional return for market risk will not increase by more than 1%

β_2: A 1% increase in SMB will result in a 0.785% increase in the additional return for market risk, keeping all else constant.

> This is plausible because a positive SMB means that the portfolio has more small cap companies in it and if a portfolio has more small-cap companies in it, it should outperform the market over the long run therefore the additional return for market risk will increase.

β_3: A 1% increase in HML will result in a 0.257% decrease in the additional return for market risk, keeping all else constant.

> This is plausible because a negative HML means that the portfolio contains a large proportion of growth stocks (Low book to market ratio), a portfolio with a large proportion of growth stocks will underperform as compared to the one with a large proportion of value stocks (High book to market ratio). Therefore, the additional return for market risk will decrease.

Hypothesis Testing:

$H_0 : \beta_1 = 0$
$H_1 : \beta_1 \neq 0$
β_1 : Since, $|t_{calc}| > t_{crit}$ (45.78 > 1.960), we reject the null hypothesis (H_0) at 5% significance level. Therefore, Market risk premium holds statistical significance at 5% level of significance.

$H_0 : \beta_2 = 0$
$H_1 : \beta_2 \neq 0$
β_2 : Since, $|t_{calc}| > t_{crit}$ (23.28 > 1.960), we reject the null hypothesis (H_0) at 5% significance level. Therefore, SMB holds statistical significance at 5% level of significance.

$H_0 : \beta_3 = 0$
$H_1 : \beta_3 \neq 0$
β_3 : Since, $|t_{calc}| > t_{crit}$ (8.51 > 1.960), we reject the null hypothesis (H_0) at 5% significance level. Therefore, HML holds statistical significance at 5% level of significance.

<u>Assessing the residuals:</u>

We need to check if there are Heteroskedastic error terms for this model. We will be using Breusch pagan test to test for Heteroskedasticity.

$H_0 : E(u_i^2) = \sigma^2$ (homoskedasticity)
$H_1 : E(u_i^2) \neq \sigma^2$ (heteroskedasticity)

We reject the null because the value of the n $\times$ $R_{\hat{u}}^2$ here is 70.16 which is larger than 7.81, the 5% critical value of the $\chi_3{}^2$ distribution. This proves that there is heteroskedasticity and we will be using HAC standard errors to get a new equation.

$$\widehat{E(R)} - R_f = 0.392 + 0.945(R_m - R_f) + 0.785\,SMB - 0.257\,HML$$
$$(0.125) \quad (0.0407) \qquad (0.0818) \quad (0.0592)$$

Even though, the estimated coefficients are the same, but the standard errors, t- statistics and p-values have changed.

We strongly reject the null hypothesis that Market risk premium, SMB and HML are insignificant when using a t statistic based on either OLS standard errors or on HAC standard errors

<u>**Industry 5: Other**</u>

CAPM

$$\widehat{E(R)} - R_f = 0.198 + 1.22(R_m - R_f)$$
$$(0.125) \quad (0.0232)$$

Interpretation of each estimated parameter:

β_0: When expected return on the market rate is equal to the risk free rate ($R_m = R_f$), the additional return for market risk will be 0.198%.

β_1: A 1% increase in the market risk premium will result in a 1.22% increase in the additional return for market risk, keeping all else constant.

> This does not make sense because the beta of the security (β_1) is greater than 1 (i.e. 1.22%) hence it is deemed to be risky. However, other industries that include services such as sanitary services will be quite stable and hence it doesn't make sense for the overall industry to be very risky. As a result, the additional return for market risk will not increase by more than 1%.

Hypothesis Testing:

$H_0 : \beta_1 = 0$
$H_1 : \beta_1 \neq 0$

β_1 : Since, $|t_{calc}| > t_{crit}$ (52.82 > 1.960), we reject the null hypothesis (H_0) at 5% significance level. Therefore, Market risk premium holds statistical significance at 5% level of significance.

Assessing the residuals:

We need to check if there are Heteroskedastic error terms for this model. We will be using Breusch pagan test to test for Heteroskedasticity.

$H_0 : E(u_i^2) = \sigma^2$ (homoskedasticity)
$H_1 : E(u_i^2) \neq \sigma^2$ (heteroskedasticity)

We reject the null because the value of the n $\times$ $R_{\hat{u}}^2$ here is 63.32 which is larger than 3.84, the 5% critical value of the χ_1^2 distribution. This proves that there is heteroskedasticity and we will be using HAC standard errors to get a new equation.

Thus the new estimated equation by using HAC standard errors is
$$\widehat{E(R)} - R_f = 0.198 + 1.22(R_m - R_f)$$

(0.119) (0.0636)

Even though, the estimated coefficients are the same, but the standard errors and t-statistics have changed.

We strongly reject the null hypothesis that Market risk premium are insignificant when using a t statistic based on either OLS standard errors or on HAC standard errors.

FAMA FRENCH

$$\widehat{E(R)} - R_f = -0.0601 + 0.961(R_m - R_f) + 0.823\,SMB + 0.685\,HML$$

(0.0651) (0.0130) (0.0213) (0.0190)

<u>Interpretation of each estimated parameter:</u>

β_0: When expected return on the market rate is equal to the risk free rate ($R_m = R_f$), the additional return for market risk will be -0.0601%.

β_1: A 1% increase in market risk premium will result in a 0.961% increase in the additional return for market risk, keeping all else constant.

> This is plausible because other industry may not only consist of services that are less risky in nature, it also might consist of services that are risky in nature. Therefore, 0.961% increase in the additional return for market risk sounds realistic.

β_2: A 1% increase in SMB will result in a 0.823% increase in the additional return for market risk, keeping all else constant.

> This is plausible because a positive SMB means that the portfolio has more small cap companies in it and if a portfolio has more small-cap companies in it, it should outperform the market over the long run therefore the additional return for market risk will increase.

β_3: A 1% increase in HML will result in a 0.685% increase in the additional return for market risk, keeping all else constant.

> This is plausible because a positive HML means that the portfolio contains a large proportion of value stocks (High book to market ratio), a portfolio with a large proportion of value stocks should outperform one with a large proportion of growth stocks (Low book to market ratio). Therefore, the additional return for market risk will increase.

<u>Hypothesis Testing:</u>

$H_0 : \beta_1 = 0$
$H_1 : \beta_1 \neq 0$
β_1 : Since, $|t_{calc}| > t_{crit}$ (73.79 > 1.960), we reject the null hypothesis (H_0) at 5% significance level. Therefore, Market risk premium holds statistical significance at 5% level of significance.

$H_0 : \beta_2 = 0$
$H_1 : \beta_2 \neq 0$
β_2 : Since, $|t_{calc}| > t_{crit}$ (38.70 > 1.960), we reject the null hypothesis (H_0) at 5% significance level. Therefore, SMB holds statistical significance at 5% level of significance.

$H_0 : \beta_3 = 0$
$H_1 : \beta_3 \neq 0$
β_3 : Since, $|t_{calc}| > t_{crit}$ (36.00 > 1.960), we reject the null hypothesis (H_0) at 5% significance level. Therefore, HML holds statistical significance at 5% level of significance.

<u>Assessing the residuals:</u>

We need to check if there are Heteroskedastic error terms for this model. We will be using Breusch pagan test to test for Heteroskedasticity.

$H_0 : E(u_i^2) = \sigma^2$ (homoskedasticity)
$H_1 : E(u_i^2) \neq \sigma^2$ (heteroskedasticity)

We reject the null because the value of the n $\times$ $R_{\hat{u}}^2$ here is 71.46 which is larger than 7.81, the 5% critical value of the χ_3^2 distribution. This proves that there is heteroskedasticity and we will be using HAC standard errors to get a new equation.

Thus, the new estimated equation after using HAC standard errors is

$$\widehat{E(R)} = R_f - \underset{(0.0676)}{0.0601} + \underset{(0.0253)}{0.961}(R_m - R_f) + \underset{(0.0590)}{0.823}\, SMB + \underset{(0.0505)}{0.685}\, HML$$

Even though, the estimated coefficients are the same, but the standard errors and t-statistics have changed.

We strongly reject the null hypothesis that Market risk premium, SMB and HML are insignificant when using a t statistic based on either OLS standard errors or on HAC standard errors

<u>Goodness of fit:</u>

In relation to the following data we are able determine which model (CAPM or FAMA French) is the stronger and more accurate model for each industry.

- Industry 1 : FAMA French is the stronger and more accurate model due to both R^2 and $\overline{R^2}$ being higher than that of CAPM (0.923, 0.923 respectively compared to 0.768, 0.767). Further, FAMA French has the lower AIC, HQ and SC (4.20, 4.21, 4.22 respectively compared to 5.30, 5.31, 5.31)
- Industry 2 : FAMA French is the stronger and more accurate model due to both R^2 and $\overline{R^2}$ being higher than that of CAPM (0.939 compared to 0.815). Further, FAMA French has the lower AIC, HQ and SC (4.01, 4.02, 4.03 respectively compared to 5.12, 5.12, 5.13)
- Industry 3 : FAMA French is the stronger and more accurate model due to both R^2 and $\overline{R^2}$ being higher than that of Industry CAPM (0.868 compared to 0.749). Further, FAMA French has the lower AIC, HQ and SC (5.07, 5.08, 5.09 respectively compared to 5.72, 5.72, 5.73)
- Industry 4 : FAMA French is the stronger and more accurate model due to both R^2 and $\overline{R^2}$ being higher than that of Industry CAPM (0.768, 0.767 respectively, compared to 0.642). Further, FAMA French has the lower AIC, HQ and SC (5.28, 5.29, 5.30 respectively compared to 5.71, 5.72, 5.72)
- Industry 5 : FAMA French is the stronger and more accurate model due to both R^2 and $\overline{R^2}$ being higher than that of Industry CAPM (0.923 compared to 0.717). Further, FAMA French has the lower AIC, HQ and SC (4.36, 4.37, 4.38 respectively compared to 5.66, 5.67, 5.67)

The overall best model is thus clearly the FAMA French Model and the strongest industry for both CAPM and FAMA French is Industry 2.

CAPM

Predictor	R^2	$\overline{R^2}$	AIC	HQ	SC
Industry 1`	0.768	0.767	5.30	5.31	5.31
Industry 2	0.815	**0.815**	**5.12**	**5.12**	**5.13**
Industry 3	0.749	0.749	5.72	5.72	5.73
Industry 4	0.642	0.642	5.71	5.72	5.72
Industry 5	0.717	0.717	5.66	5.67	5.67

FAMA FRENCH

Predictor	R^2	$\overline{R^2}$	AIC	HQ	SC
Industry 1	0.923	0.923	4.20	4.21	4.22
Industry 2	0.939	**0.939**	**4.01**	**4.02**	**4.03**

Industry 3	0.868	0.868	5.07	5.08	5.09
Industry 4	0.768	0.767	5.28	5.29	5.30
Industry 5	0.923	0.923	4.36	4.37	4.38

In relation to the two models, Industry 2 is the best.

PART B

(i) Small Firms

When analysing small firms, the "lo30" portfolio was selected. The value-weighted returns were subtracted by the risk-free rate to calculate excess returns. January was used as the base variable and dummy variables were created for the months of February to December.

$$\begin{aligned} &Return_t \\ &= \beta_0 + \beta_1 Feb_t + \beta_2 Mar_t + \beta_3 Apr_t + \beta_4 May_t + \beta_5 Jun_t + \beta_6 Jul_t + \beta_7 Aug_t + \beta_8 Sept_t + \beta_9 Oct_t + \beta_{10} Nov_t + \beta_{11} Dec_t \\ &= \underset{(0.860)}{4.68} - \underset{(1.217)}{3.44} Feb_t - \underset{(1.217)}{4.01} Mar_t - \underset{(1.220)}{3.64} Apr_t - \underset{(1.220)}{3.99} May_t - \underset{(1.220)}{4.03} Jun_t - \underset{(1.217)}{3.38} Jul_t - \underset{(1.217)}{3.78} Aug_t - \underset{(1.217)}{5.34} Sept_t - \underset{(1.217)}{5.55} Oct_t - \underset{(1.217)}{3.46} Nov_t - \underset{(1.217)}{3.64} Dec_t \end{aligned}$$

Assessing the Residuals:

Heteroskedasticity:

If model is affected by heteroskedasticity, the standard errors and therefore t-statistics will no longer be reliable. The Breusch-Pagan test will be used to test for heteroskedastic errors.

$H_0 : E(u_i^2) = \sigma^2$ (homoskedasticity)
$H_1 : E(u_i^2) \neq \sigma^2$ (heteroskedasticity)

As $F_{calc} < F_{crit}$ (0.6115 < 1.80), at 5% significance we cannot reject the null hypothesis (H_0) and conclude that there is insufficient evidence that the model is heteroskedastic.

<u>Autocorrelation:</u>
If model is affected by autocorrelation, the standard errors and therefore t-statistics will no longer be reliable. The Breusch-Godfrey test will be used to test for autocorrelation in the model.

H_0: $\rho_1 = \rho_1 = \rho_2 = \rho_3 = \rho_4 = \rho_5 = \rho_6 = \rho_7 = \rho_8 = \rho_9 = \rho_{10} = \rho_{11} = \rho_{12} = 0$
H_0: $\rho_i \neq 0$

As BG_{calc} (87.59) is larger BG_{crit} at 5% significance (21.03), at 5% significance level, we can reject the null hypothesis (H_0) and conclude that there is evidence of serial correlation.

<u>Adjusting for Autocorrelation:</u>
As we found evidence of serial correlation, the standard errors must be adjusted. We do this using HAC standard errors.

The new estimated equation based on HAC errors:

$$\widehat{Return}_t = \underset{(0.825)}{4.68} - \underset{(0.879)}{3.44}Feb_t - \underset{(1.049)}{4.01}Mar_t - \underset{(1.272)}{3.64}Apr_t - \underset{(1.454)}{3.99}May_t - \underset{(1.110)}{4.03}Jun_t - \underset{(1.227)}{3.38}Jul_t - \underset{(1.359)}{3.78}Aug_t - \underset{(1.332)}{5.34}Sept_t - \underset{(1.193)}{5.55}Oct_t - \underset{(1.119)}{3.46}Nov_t - \underset{(1.060)}{3.64}Dec_t$$

Interpretation of Parameters

β_0 : The portfolio is expected on average to generate 4.68% excess return in January holding all else constant.

This is plausible as it is expected that the portfolio would generate a positive return in January.

β_1 : The portfolio is expected on average to generate 3.44% less excess return in February compared to January holding all else constant.

This is plausible that February would generate less return than January due to the "January effect" theory which details that higher returns are observed in January than any other month.

β_2 : The portfolio is expected on average to generate 4.01% less excess return in March compared to January holding all else constant.

This is plausible that March would generate less return than January due to the "January effect" theory which details that higher returns are observed in January than any other month.

β_3 : The portfolio is expected on average to generate 3.64% less excess return in April compared to January holding all else constant.

This is plausible that April would generate less return than January due to the "January effect" theory which details that higher returns are observed in January than any other month.

β_4 : The portfolio is expected on average to generate 3.99% less excess return in May compared to January holding all else constant.

This is plausible that May would generate less return than January due to the "January effect" theory which details that higher returns are observed in January than any other month.

β_5 : The portfolio is expected on average to generate 4.03% less excess return in June compared to January holding all else constant.

This is plausible that June would generate less return than January due to the "January effect" theory which details that higher returns are observed in January than any other month.

β_6 : The portfolio is expected on average to generate 3.38% less excess return in July compared to January holding all else constant.

This is plausible that July would generate less return than January due to the "January effect" theory which details that higher returns are observed in January than any other month.

β_7 : The portfolio is expected on average to generate 3.78% less excess return in August compared to January holding all else constant.

This is plausible that August would generate less return than January due to the "January effect" theory which details that higher returns are observed in January than any other month.

β_8 : The portfolio is expected on average to generate 5.34% less excess return in September compared to January holding all else constant.

This is plausible that September would generate less return than January due to the "January effect" theory which details that higher returns are observed in January than any other month.

β_9 : The portfolio is expected on average to generate 5.55% less excess return in October compared to January holding all else constant.

This is plausible that October would generate less return than January due to the "January effect" theory which details that higher returns are observed in January than any other month.

β_{10} : The portfolio is expected on average to generate 3.46% less excess return in November compared to November holding all else constant.

This is plausible that February would generate less return than January due to the "January effect" theory which details that higher returns are observed in January than any other month.

β_{11} : The portfolio is expected on average to generate 3.64% less excess return in December compared to January holding all else constant.

This is plausible that December would generate less return than January due to the "January effect" theory which details that higher returns are observed in January than any other month.

Hypothesis Testing:

T-tests to determine if the variables are statistically significant.

$H_0 : \beta_1 = 0$
$H_1 : \beta_1 \neq 0$
β_1 : As $|t_{calc}| > t_{crit}$ (3.92 > 1.960), we reject the null hypothesis at 5% significance and conclude that the February variable is statistically significant.

$H_0 : \beta_2 = 0$
$H_1 : \beta_2 \neq 0$
β_2 : As $|t_{calc}| > t_{crit}$ (3.82 > 1.960), we reject the null hypothesis at 5% significance and conclude that the March variable is statistically significant.

$H_0 : \beta_3 = 0$
$H_1 : \beta_3 \neq 0$
β_3 : As $|t_{calc}| > t_{crit}$ (2.86 > 1.960), we reject the null hypothesis at 5% significance and conclude that the April variable is statistically significant.

$H_0 : \beta_4 = 0$
$H_1 : \beta_4 \neq 0$
β_4 : As $|t_{calc}| > t_{crit}$ (2.75 > 1.960), we reject the null hypothesis at 5% significance and conclude that the May variable is statistically significant.

$H_0 : \beta_5 = 0$
$H_1 : \beta_5 \neq 0$
β_5 : As $|t_{calc}| > t_{crit}$ (3.63 > 1.960), we reject the null hypothesis at 5% significance and conclude that the June variable is statistically significant.

$H_0 : \beta_6 = 0$
$H_1 : \beta_6 \neq 0$
β_6 : As $|t_{calc}| > t_{crit}$ (2.75 > 1.960), we reject the null hypothesis at 5% significance and conclude that the July variable is statistically significant.

$H_0 : \beta_7 = 0$
$H_1 : \beta_7 \neq 0$
β_7 : As $|t_{calc}| > t_{crit}$ (2.77 > 1.960), we reject the null hypothesis at 5% significance and conclude that the August variable is statistically significant.

$H_0 : \beta_8 = 0$
$H_1 : \beta_8 \neq 0$
β_8 : As $|t_{calc}| > t_{crit}$ (4.01 > 1.960), we reject the null hypothesis at 5% significance and conclude that the September variable is statistically significant.

$H_0 : \beta_9 = 0$
$H_1 : \beta_9 \neq 0$
β_9 : As $|t_{calc}| > t_{crit}$ (4.65 > 1.960), we reject the null hypothesis at 5% significance and conclude that the October variable is statistically significant.

$H_0 : \beta_{10} = 0$
$H_1 : \beta_{10} \neq 0$
β_{10} : As $|t_{calc}| > t_{crit}$ (3.09 > 1.960), we reject the null hypothesis at 5% significance and conclude that the November variable is statistically significant.

$H_0 : \beta_{11} = 0$
$H_1 : \beta_{11} \neq 0$
β_{11} : As $|t_{calc}| > t_{crit}$ (3.43 > 1.960), we reject the null hypothesis at 5% significance and conclude that the December variable is statistically significant.

As all the variables are statistically significant and generate less excess returns than in January, there is evidence of the "January effect".

(ii) Large Firms

When analysing large firms, the "hi30" portfolio was selected. The value-weighted returns were subtracted by the risk-free rate to calculate excess returns. January was used as the base variable and dummy variables were created for the months of February to December.

$$Return_t = \beta_0 + \beta_1 Feb_t + \beta_2 Mar_t + \beta_3 Apr_t + \beta_4 May_t + \beta_5 Jun_t + \beta_6 Jul_t + \beta_7 Aug_t + \beta_8 Sept_t + \beta_9 Oct_t + \beta_{10} Nov_t + \beta_{11} Dec_t$$

$$= \underset{(0.54)}{0.90} - \underset{(0.762)}{0.61} Feb_t - \underset{(0.762)}{0.39} Mar_t + \underset{(0.764)}{0.17} Apr_t - \underset{(0.764)}{0.65} May_t - \underset{(0.764)}{0.31} Jun_t + \underset{(0.762)}{0.40} Jul_t - \underset{(0.762)}{0.05} Aug_t - \underset{(0.762)}{1.89} Sept_t - \underset{(0.762)}{0.60} Oct_t + \underset{(0.762)}{0.37} Nov_t + \underset{(0.762)}{0.49} Dec_t$$

Assessing the Residuals:

<u>Testing for Heteroskedasticity:</u>
If model is affected by heteroskedasticity, the standard errors and therefore t-statistics will no longer be reliable. We use the Breusch-Pagan test to test for heteroskedastic errors

$H_0 : E(u_i^2) = \sigma^2$ (homoskedasticity)
$H_1 : E(u_i^2) \neq \sigma^2$ (heteroskedasticity)

As $F_{calc} < F_{crit}$ (0.828 < 1.80), at 5% significance we cannot reject the null hypothesis (H_0) that the model is homoskedastic and conclude that there is insufficient evidence that the model is heteroskedastic.

Autocorrelation:
If model is affected by autocorrelation, the standard errors and therefore t-statistics will no longer be reliable. We test for autocorrelation in the model using the Breusch-Godfrey test.

H_0: $\rho_1 = \rho_1 = \rho_2 = \rho_3 = \rho_4 = \rho_5 = \rho_6 = \rho_7 = \rho_8 = \rho_9 = \rho_{10} = \rho_{11} = \rho_{12} = 0$
H_0: $\rho_i \neq 0$

As BG_{calc} (35.25) is greater than BG_{crit} at 5% significance (21.03), at 5% significance level, we can reject the null hypothesis (H_0) and conclude that there is evidence of serial correlation.

Adjusting for Autocorrelation:
As we found evidence of serial correlation, the standard errors must be adjusted. We do this using HAC standard errors.

The new estimated equation based on HAC errors:

$$\begin{aligned}
&Return_t\\
&= \underset{(0.479)}{0.90} - \underset{(0.595)}{0.61}Feb_t - \underset{(0.687)}{0.39}Mar_t + \underset{(0.788)}{0.17}Apr_t - \underset{(0.738)}{0.65}May_t - \underset{(0.691)}{0.31}Jun_t + \underset{(0.770)}{0.40}Jul_t - \underset{(0.760)}{0.05}Aug_t - \underset{(0.773)}{1.89}Sept_t - \underset{(0.771)}{0.60}Oct_t + \underset{(0.684)}{0.37}Nov_t + \underset{(0.604)}{0.49}Dec_t
\end{aligned}$$

Interpretation of Parameters

β_0 : The portfolio is expected on average to generate 0.90% excess return in January holding all else constant.

> This is plausible for a portfolio with large firms to have a small positive return in January.

β_1 : The portfolio is expected on average to generate 0.61% less excess return in February compared to January holding all else constant.

> This is plausible that the portfolio generates slightly less return in February than January as it is expected that excess returns would vary slightly between each month.

β_2 : The portfolio is expected on average to generate 0.39% less excess return in March compared to January holding all else constant.

> This is plausible that the portfolio generates slightly less return in March than January as it is expected that excess returns would vary slightly between each month.

β_3 : The portfolio is expected on average to generate 0.17% more excess return in April compared to January holding all else constant.

> This is plausible that the portfolio generates slightly more return in April than January as it is expected that excess returns would vary slightly between each month.

β_4 : The portfolio is expected on average to generate 0.65% less excess return in May compared to January holding all else constant.

> This is plausible that the portfolio generates slightly less return in May than January as it is expected that excess returns would vary slightly between each month.

β_5 : The portfolio is expected on average to generate 0.31% less excess return in June compared to January holding all else constant.

> This is plausible that the portfolio generates slightly less return in June than January as it is expected that excess returns would vary slightly between each month.

β_6 : The portfolio is expected on average to generate 0.40% more excess return in July compared to January holding all else constant.

> This is plausible that the portfolio generates slightly more return in July than January as it is expected that excess returns would vary slightly between each month.

β_7 : The portfolio is expected on average to generate 0.05% less excess return in August compared to January holding all else constant.

> This is plausible that the portfolio generates slightly less return in August than January as it is expected that excess returns would vary slightly between each month.

β_8 : The portfolio is expected on average to generate 1.89% less excess return in September compared to January holding all else constant.

> This is plausible that the portfolio generates slightly less return in September than January as it is expected that excess returns would vary slightly between each month.

β_9 : The portfolio is expected on average to generate 0.60% less excess return in October compared to January holding all else constant.

> This is plausible that the portfolio generates slightly less return in October than January as it is expected that excess returns would vary slightly between each month.

β_{10} : The portfolio is expected on average to generate 0.37% more excess return in November compared to January holding all else constant.

> This is plausible that the portfolio generates slightly more return in November than January as it is expected that excess returns would vary slightly between each month.

β_{11} : The portfolio is expected on average to generate 0.49% more excess return in December compared to January holding all else constant.

> This is plausible that the portfolio generates more less return in December than January as it is expected that excess returns would vary slightly between each month.

<u>Hypothesis Testing:</u>

T-tests to determine if the variables are statistically significant.

$H_0 : \beta_1 = 0$
$H_1 : \beta_1 \neq 0$
β_1 : As $|t_{calc}| < t_{crit}$ (1.02 < 1.960), we cannot reject the null hypothesis at 5% significance and conclude that the February variable is not statistically significant.

$H_0 : \beta_2 = 0$
$H_1 : \beta_2 \neq 0$
β_2 : As $|t_{calc}| > t_{crit}$ (0.56 < 1.960), we cannot reject the null hypothesis at 5% significance and conclude that the March variable is not statistically significant.

$H_0 : \beta_3 = 0$
$H_1 : \beta_3 \neq 0$
β_3 : As $|t_{calc}| > t_{crit}$ (0.21 < 1.960), we cannot reject the null hypothesis at 5% significance and conclude that the April variable is not statistically significant.

$H_0 : \beta_4 = 0$
$H_1 : \beta_4 \neq 0$
β_4 : As $|t_{calc}| > t_{crit}$ (0.88 < 1.960), we cannot reject the null hypothesis at 5% significance and conclude that the May variable is not statistically significant.

$H_0 : \beta_5 = 0$
$H_1 : \beta_5 \neq 0$
β_5 : As $|t_{calc}| > t_{crit}$ (0.45 < 1.960), we cannot reject the null hypothesis at 5% significance and conclude that the June variable is statistically significant.

$H_0 : \beta_6 = 0$
$H_1 : \beta_6 \neq 0$
β_6 : As $|t_{calc}| > t_{crit}$ (0.52 < 1.960), we cannot reject the null hypothesis at 5% significance and conclude that the July variable is statistically significant.

$H_0 : \beta_7 = 0$
$H_1 : \beta_7 \neq 0$
β_7 : As $|t_{calc}| > t_{crit}$ (0.07 < 1.960), we cannot reject the null hypothesis at 5% significance and conclude that the August variable is statistically significant.

$H_0 : \beta_8 = 0$
$H_1 : \beta_8 \neq 0$
β_8 : As $|t_{calc}| > t_{crit}$ (2.44 > 1.960), we reject the null hypothesis at 5% significance and conclude that the September variable is statistically significant.

$H_0 : \beta_9 = 0$
$H_1 : \beta_9 \neq 0$
β_9 : As $|t_{calc}| > t_{crit}$ (0.77 < 1.960), we cannot reject the null hypothesis at 5% significance and conclude that the October variable is not statistically significant.

$H_0 : \beta_{10} = 0$
$H_1 : \beta_{10} \neq 0$
β_{10} : As $|t_{calc}| > t_{crit}$ (0.54 < 1.960), we cannot reject the null hypothesis at 5% significance and conclude that the November variable is not statistically significant.

$H_0 : \beta_{11} = 0$
$H_1 : \beta_{11} \neq 0$
β_{11} : As $|t_{calc}| > t_{crit}$ (0.81 < 1.960), we cannot reject the null hypothesis at 5% significance and conclude that the December variable is not statistically significant.

Only the september variable is statistically significant and there is no evidence of “January effect” in large firms

Comparison
After completing the regression and observing the models with the small firms and large firms, the “January effect” is only observed in the small firms. This could be because large firms are not as volatile and therefore do not contain evidence of this.

Goodness of Fit:

Predictor	R^2	$\overline{R^2}$	AIC	HQ	SC
Small Firms	0.246	0.0147	7.07	7.09	7.12
Large Firms	0.0148	0.0049	6.133	6.153	6.187

In relation to the following data we are able determine whether there is any seasonal “January effect” in stock market by comparing small and large firms.

- Smaller firms demonstrate a stronger and more accurate seasonal “January effect” due to both R^2 and $\overline{R^2}$ being higher than that of larger firms (0.246,0.0147 respectively compared to 0.0148,0.0049).
- Furthermore, AIC for small firms is larger (7.07) compared to larger firms (6.133), and HQ ad SC are also are larger for smaller firms (7.09, 7.12) compared to larger firms (6.153, 6.187).

Therefore from given data, we cannot conclude anything as the data is contradicting. Small firms would best fit according to R^2 and $\overline{R^2}$ and according to other measurement variables larger firms would be more suited.

Conclusion and Discussion

According to the research that was completed, it was suggested that the Fama and French 3 Factor Model would provide more conclusive and accurate predictions about the expected return and risk of the industry portfolio's at hand. This is due to the extra variables (size and value factors) that are a part of the model (an extension of CAPM). According to the results that were gained from the regression outputs, it can be seen that for each industry, the following asset pricing models provide the strongest results due to the hypothesis testing and the goodness of fit and predictive power which was performed.

Based on the results of Part A, in regards to Industry 1; FAMA French is the stronger and more accurate model due to both R^2 and $\overline{R^2}$ being higher than that of CAPM (0.923 compared to 0.768 and 7.67 respectively). Further, FAMA French has the lower AIC, HQ and SC (4.20, 4.21, 4.22 respectively compared to 5.30, 5.31, 5.31). Further, in Industry 2; FAMA French is the stronger and more accurate model due to both R^2 and $\overline{R^2}$ being higher than that of CAPM (0.939 compared to 0.815). Further, FAMA French has the lower AIC, HQ and SC (4.01, 4.02, 4.03 respectively compared to 5.12, 5.12, 5.13). The results regarding Industry 3, highlight that FAMA French is the stronger and more accurate model due to both R^2 and $\overline{R^2}$ being higher than that of Industry CAPM (0.868 compared to 0.749). Further, FAMA French has the lower AIC, HQ and SC (5.07, 5.08, 5.09 respectively compared to 5.72, 5.72, 5.73). Industry 4 further follows the pattern of the previous industries with FAMA French is the stronger and more accurate model due to both R^2 and $\overline{R^2}$ being higher than that of Industry CAPM (0.768 and 0.767 respectively, compared to 0.642). Further, FAMA

French has the lower AIC, HQ and SC (5.28, 5.29, 5.30 respectively compared to 5.71, 5.72, 5.72) Lastly, Industry 5 also supports the idea that FAMA French is the stronger and more accurate model due to both R^2 and $\overline{R^2}$ being higher than that of Industry CAPM (0.923 compared to 0.717). Further, FAMA French has the lower AIC, HQ and SC (4.36, 4.37, 4.38 respectively compared to 5.66, 5.67, 5.67)

Further all of the models contain heteroskedasticity and therefore the HACC was utilised in order to come up with a new equation. Thus, while the coefficients of the parameters did not change, the standard errors and t statistic of the parameters changed. As a result of these changes, the changes to the t statistic may thus affect whether or not the variables are statistically significant.

Thus, as predicted, the Fama and French 3 factor model is the most effective compared to the CAPM model and the added variables do indeed provide greater results.

In conclusion, for part B, we can say “January effect” was present in small firms because te portfolio return rate from march to december are lower than January as shown in Model 2. However the model does not have a goodness of fit level to completely support the theory. The goodness of fit is contradictory and makes it hard to choose one between small and large firms. Few factors that can improve the model are differences between market rate and risk free rates and most recent quality data. Also, since the R^2 value is small (0.0049) for large firms, it demonstrates that only .49% of data is sown by the model and does not give accurate and enough information of larger firms’ portfolios. After completing the regression and observing the models with the small firms and large firms, the “January effect” is only observed in the small firms. This could be because large firms are not as volatile and therefore do not contain evidence of this.

Thus, based on results of regression and hypothesis testing, small firms demonstrate the “January effect”.

While it is never going to be possible to fully predict the outcome of stocks and portfolios, changes can be made to improve the result. The 3 factor model is unable to explain the poor performance of small growth stocks, and this could be improved by including profitability and investment variables, which are included in the 5 factor model. Further, the Fama and French 3 factor model does not include a variable for momentum (the tendency for the stock price to continue rising if it is going up and continue declining if it is going down). Similarly, other factors could have been included in the model is the low volatility factor (which includes stocks that have low variation measured by statistical measures including variance or beta values). This would help to include new variable data that was not previously considered, thus hopefully increasing accuracy for those who wish to use the models to predict the future of their returns.

In order to improve the performance of the model; they could utilise past information to help predict future outcomes. While continuing to look at stock prices; they could also consider the firm’s book to market ratio and its return on equity. Book-to-market refers to a company's book value divided by its market capitalization and provides greater insight into predictability. Return on equity assesses how efficiently a company generates profits with the money gained from shareholders. These two methods will help contribute to the current CAPM method and assist in providing greater insight into future trends within the market.