Solution to Exercise E6.

The Multiple Regression Model. Inference

Exercise E6.1 Beach umbrella rental

Part I. Simple Linear Regression Model.

a. Regression model: $U_t = \alpha + \beta T_t + u_t \quad t = 1, \ldots, 22$

Model 1: OLS, using observations 2013-05-05–2013-09-29 ($T = 22$)
Dependent variable: U

	Coefficient	Std. Error	t-ratio	p-value
const	27.0692	26.9160	1.0057	0.3266
T	11.4595	0.860078	13.3238	0.0000

Mean dependent var	381.2727	S.D. dependent var	60.60110
Sum squared resid	7808.908	S.E. of regression	19.75969
R^2	0.898747	Adjusted R^2	0.893684
$F(1, 20)$	177.5241	P-value(F)	2.09e–11
Log-likelihood	−95.80841	Akaike criterion	195.6168
Schwarz criterion	197.7989	Hannan–Quinn	196.1308
$\hat{\rho}$	0.127938	Durbin–Watson	1.661439

SRF: $\hat{S}_t = 27.0692 + 11.4595\, T_t$

b. Test the statistical significance of the variable temperature.

$$H_0 : \beta = 0 \qquad H_a : \beta \neq 0 \qquad t = \frac{\hat{\beta} - 0}{\hat{\sigma}_{\hat{\beta}}} \overset{H_0}{\sim} t(T-k)$$

Given that $|\, t \,| = 13.3238 > 2.08596 = t_{0.025}(20)$, the null hypothesis is rejected at the 5% significance level. Therefore, the temperature is a statistically significant variable.

c. Two-sided test.

$$H_0 : \beta = 20/2 = 10 \qquad H_a : \beta \neq 20/2 \qquad t = \frac{\hat{\beta} - 10}{\hat{\sigma}_{\hat{\beta}}} \overset{H_0}{\sim} t(T-k)$$

Since $|\, t \,| = \left| \frac{11.4595 - 10}{0.860078} \right| = |\, 1.6969 \,| < 2.08596 = t_{0.025}(20)$, the null hypothesis is not rejected at the 5% significance level. Therefore, if the average temperature increases by 2°C, the number of rented umbrellas may increase by 20 units.

d. Point prediction.

$\hat{U}_p = 27.0692 + 11.4595 \times 42 = 508.3682$ umbrellas.

e. Prediction interval.

$CI(U_p)_{0.95} = [461.76\,;\, 554.97]$

The estimated upper limit for the number of rented umbrellas is 554.97.

Part II. General Linear Regression Model.

a. Regression model: $U_t = \gamma_1 + \gamma_2 \, T_t + \gamma_3 \, P_t + \gamma_4 \, WW_t + v_t \quad t = 1, \ldots, 22$

Model 3: OLS, using observations 2013-05-05–2013-09-29 ($T = 22$)
Dependent variable: U

	Coefficient	Std. Error	t-ratio	p-value
const	44.5693	45.7575	0.9740	0.3429
T	11.0471	2.18129	5.0645	0.0001
P	−0.0524909	3.32631	−0.0158	0.9876
WW	−10.5245	9.73983	−1.0806	0.2942

Mean dependent var	381.2727	S.D. dependent var	60.60110
Sum squared resid	7320.861	S.E. of regression	20.16716
R^2	0.905075	Adjusted R^2	0.889254
$F(3, 18)$	57.20762	P-value(F)	2.11e–09
Log-likelihood	−95.09850	Akaike criterion	198.1970
Schwarz criterion	202.5612	Hannan–Quinn	199.2251
$\hat{\rho}$	0.107465	Durbin–Watson	1.685968

SRF: $\hat{U}_t = 44.5693 + 11.0471 \, T_t - 0.0524909 \, P_t - 10.5245 \, WW_t \quad t = 1, \ldots, 22$

b. Test whether the variable temperature is statistically significant.

$$H_0 : \gamma_2 = 0 \qquad H_a : \gamma_2 \neq 0 \qquad t = \frac{\hat{\gamma}_2 - 0}{\hat{\sigma}_{\hat{\gamma}_2}} \overset{H_0}{\sim} t(T-k)$$

Since $| \, t \, | = 5.0645 > 2.10092 = t_{0.025}(18)$, the null hypothesis is rejected at the 5% significance level. Therefore, the variable temperature is statistically significant.

Test whether the variable price is statistically significant.

$$H_0 : \gamma_3 = 0 \qquad H_a : \gamma_3 \neq 0 \qquad t = \frac{\hat{\gamma}_3 - 0}{\hat{\sigma}_{\hat{\gamma}_3}} \overset{H_0}{\sim} t(T-k)$$

Since $| \, t \, | = 0.0158 < 2.10092 = t_{0.025}(18)$, the null hypothesis is not rejected at the 5% significance level. Therefore, the variable price is not statistically significant.

Test whether the variable *windy week* is statistically significant.

$$H_0 : \gamma_4 = 0 \qquad H_a : \gamma_4 \neq 0 \qquad t = \frac{\hat{\gamma}_4 - 0}{\hat{\sigma}_{\hat{\gamma}_4}} \overset{H_0}{\sim} t(T-k)$$

Given that $| \, t \, | = 1.0806 < 2.10092 = t_{0.025}(18)$, the null hypothesis is not rejected at the 5% significance level. Therefore, the variable windy week is not statistically significant.

Test the overall significance of the explanatory variables.

$$H_0 : \gamma_2 = \gamma_3 = \gamma_4 = 0 \qquad H_a : \exists \gamma_i \neq 0 \qquad F = \frac{R^2/(k-1)}{(1-R^2)/(T-k)} \overset{H_0}{\sim} \mathcal{F}(k-1, T-k)$$

Since $F = 57.20762 > 3.15991 = \mathcal{F}_{0.05}(3, 18)$, the null hypothesis is rejected at the 5% significance level. Therefore, the explanatory variables are jointly significant.

c. Given that the variables price and *windy week* are not individually significant, it may be possible that there is a high degree of collinearity between these two variables. Let's test whether the variables price and *windy week* are jointly significant.

$$H_0 : \gamma_3 = \gamma_4 = 0 \qquad F = \frac{SSR_R - SSR_{UR}}{SSR_{UR}} \frac{T-k}{q} \overset{H_0}{\sim} \mathcal{F}(q, T-k)$$
$$H_a : \exists \gamma_i \neq 0$$

Given that $F = 0.599988 < 3.55456 = \mathcal{F}_{0.05}(2, 18)$, the null hypothesis is not rejected at the 5% significance level. Therefore, variables price and *windy week* are not jointly significant.

The variables price and *windy week* are neither individually nor jointly significant. Therefore, it may be concluded that there is not a high degree of collinearity in the sample.

d. Prediction interval.

Windy week: $CI(U_p)_{0.95} = [407.05\,;\,544.02]$

The estimated upper limit for the number of rented umbrellas would be 544.02.

Non windy week: $CI(U_p)_{0.95} = [422.60\,;\,549.53]$

The estimated upper limit for the number of rented umbrellas would be 549.53.

e. Model (2). The variables price and *windy week* are neither individually nor jointly significant. Therefore, the inclusion of the true restrictions $\beta_3 = \beta_4 = 0$ in the estimation of the model reduces the variance of the estimator.

Exercise E6.2 Holiday cottages

Model A

Regression model: $RP_i = \alpha_1 + \alpha_2\, NR_i + \alpha_3\, BP_i + u_i \quad i = 1, 2, \ldots, 75$

Model 1: OLS, using observations 1–75
Dependent variable: RP

	Coefficient	Std. Error	t-ratio	p-value
const	38.4321	7.22899	5.3164	0.0000
NR	2.26766	1.20082	1.8884	0.0630
BP	1.49558	1.09746	1.3628	0.1772

Mean dependent var	56.13893	S.D. dependent var	14.98446
Sum squared resid	15263.15	S.E. of regression	14.55982
R^2	0.081392	Adjusted R^2	0.055875
$F(2, 72)$	3.189724	P-value(F)	0.047064
Log-likelihood	−305.7595	Akaike criterion	617.5189
Schwarz criterion	624.4714	Hannan–Quinn	620.2950

SRF: $\widehat{RP}_i = 38.4321 + 2.26766\, NR_i + 1.49558\, BP_i$

a. Test the joint significance of the explanatory variables.

$$H_0 : \alpha_2 = \alpha_3 = 0 \qquad H_a : \exists \alpha_i \neq 0 \qquad F = \frac{R^2/(k-1)}{(1-R^2)/(T-k)} \overset{H_0}{\sim} \mathcal{F}(k-1, T-k)$$

Since $F = 3.189724 > 3.12391 = \mathcal{F}_{0.05}(2, 72)$, the null hypothesis is rejected at the 5% significance level. Therefore, the explanatory variables are jointly significant.

b. Test the statistically significance of the variable number of rooms.

$$H_0 : \alpha_2 = 0 \qquad H_a : \alpha_2 \neq 0 \qquad t = \frac{\hat{\alpha}_2 - 0}{\hat{\sigma}_{\hat{\alpha}_2}} \overset{H_0}{\sim} t(T-k)$$

Since $\mid t \mid = 1.888 < 1.99346 = t_{0.025}(72)$, the null hypothesis is not rejected at the 5% significance level. Therefore, the variable number of rooms is not statistically significant.

c. Confidence interval for the coefficient α_3.

$CI(\alpha_3)_{0.95} = [-0.692164\,;\, 3.68333]$

If the price of breakfast increases by €1, holding the number of rooms fixed, the variation in the price of the room lies between -0.692164 and 3.68333 euros, with a 95% probability.

d. Point prediction.

$\widehat{RP}_p = 38.4321 + 2.26766 \times 10 + 1.49558 \times 3 = 65.59544$ euros.

e. Prediction interval.

$CI(RP_p)_{0.95} = [33.345\,;\, 97.846]$

If a holiday cottage has 10 bedrooms and the price of breakfast is €3, the price of the room lies between 33.345 and 97.846 euros, with a 95% probability.

Model B

a. Regression model:

$$RP_i = \lambda_1 + \lambda_2\, NR_i + \lambda_3\, BP_i + \lambda_4\, WIFIF_i + \lambda_5\, WIFIP_i + \lambda_6\, LOCC_i + u_i$$

Model 2: OLS, using observations 1–75
Dependent variable: RP

	Coefficient	Std. Error	t-ratio	p-value
const	40.5761	7.39661	5.4858	0.0000
NR	1.94192	1.21303	1.6009	0.1140
BP	0.559911	1.21918	0.4593	0.6475
WIFIF	6.98544	3.65362	1.9119	0.0600
WIFIP	−5.75696	12.0827	−0.4765	0.6352
LOCC	2.11170	5.43209	0.3887	0.6987

Mean dependent var	56.13893	S.D. dependent var	14.98446
Sum squared resid	14429.98	S.E. of regression	14.46133
R^2	0.131536	Adjusted R^2	0.068604
$F(5, 69)$	2.090130	P-value(F)	0.077001
Log-likelihood	−303.6545	Akaike criterion	619.3089
Schwarz criterion	633.2138	Hannan–Quinn	624.8610

SRF:

$$\begin{aligned}\widehat{RP}_i &= 40.5761 + 1.94192\, NR_i + 0.559911\, BP_i + 6.98544\, WIFIF_i - \\ &- 5.75696\, WIFIP_i + 2.11170\, LOCC_i\end{aligned}$$

b. Test whether the variables *access to WiFi* and *location* are jointly significant.

$$\begin{array}{l} H_0 : \lambda_4 = \lambda_5 = \lambda_6 = 0 \\ H_a : \exists \lambda_i \neq 0 \end{array} \qquad F = \frac{SSR_R - SSR_{UR}}{SSR_{UR}} \frac{T-k}{q} \overset{H_0}{\sim} \mathcal{F}(q, T-k)$$

Since $F = 1.328 < 2.73749 = \mathcal{F}_{0.05}(3, 69)$, the null hypothesis is not rejected at the 5% significance level. Therefore, the variables *access to WiFi* and *location* are not jointly significant.

c. One-sided test.

$$\begin{array}{l} H_0 : \lambda_6 \leq 0 \\ H_a : \lambda_6 > 0 \end{array} \qquad t = \frac{\hat{\alpha}_6 - 0}{\hat{\sigma}_{\hat{\alpha}_6}} \overset{H_0}{\sim} t(T-k)$$

Since $t = 0.3887 < 1.66724 = t_{0.05}(69)$, the null hypothesis is not rejected at the 5% significance level. Therefore, holiday cottages are not more expensive in the town center.

d. Test whether variable *access to WiFi* is statistically significant.

$$\begin{array}{l} H_0 : \lambda_4 = \lambda_5 = 0 \\ H_a : \exists \lambda_i \neq 0 \end{array} \qquad F = \frac{SSR_R - SSR_{UR}}{SSR_{UR}} \frac{T-k}{q} \overset{H_0}{\sim} \mathcal{F}(q, T-k)$$

Given that $F = 1.98942 < 3.12964 = \mathcal{F}_{0.05}(2, 69)$, the null hypothesis is not rejected at the 5% significance level. Therefore, the explanatory variable *access to WiFi* is not statistically significant.

e. Test for the equality of coefficients.

$$H_0 : \lambda_4 = \lambda_5 \qquad H_a : \lambda_4 \neq \lambda_5 \qquad F = \frac{SSR_R - SSR_{UR}}{SSR_{UR}} \frac{T-k}{q} \overset{H_0}{\sim} \mathcal{F}(q, T-k)$$

Since $F = 1.04678 < 3.97981 = \mathcal{F}_{0.05}(1, 69)$, the null hypothesis is not rejected at the 5% significance level. Therefore, there is evidence in the sample in favor of the hypothesis that what is really important is to offer access to WiFi; it does not matter whether it is free or not.

Model C

a. $RP_i = \beta_1 + \beta_2\, NR_i + \beta_3\, BP_i + \beta_4\, WIFIF_i + \beta_5\, NPR_i + \beta_6\, BER_i + \beta_7\, LKR_i + u_i$

Model 3: OLS, using observations 1–75
Dependent variable: RP

	Coefficient	Std. Error	t-ratio	p-value
const	33.9408	6.72069	5.0502	0.0000
NR	1.36156	1.11199	1.2244	0.2250
BP	1.62738	1.07599	1.5124	0.1351
WIFIF	9.02246	3.32099	2.7168	0.0084
NPR	3.33934	3.93959	0.8476	0.3996
BER	16.1587	4.47207	3.6133	0.0006
LKR	12.0185	7.81149	1.5386	0.1285

Mean dependent var	56.13893	S.D. dependent var	14.98446
Sum squared resid	11550.44	S.E. of regression	13.03301
R^2	0.304840	Adjusted R^2	0.243503
$F(6, 68)$	4.969876	P-value(F)	0.000285
Log-likelihood	−295.3075	Akaike criterion	604.6151
Schwarz criterion	620.8375	Hannan–Quinn	611.0925

SRF:

$$\begin{aligned}\widehat{RP}_i &= 33.9408 + 1.36156\, NR_i + 1.62738\, BP_i + 9.02246\, WIFIF_i + \\ &+ 3.33934\, NPR_i + 16.1587\, BER_i + 12.0185\, LKR_i\end{aligned}$$

b. Test the statistical significance of the variable *having free access to WiFi.*

$$H_0 : \beta_4 = 0 \qquad H_a : \beta_4 \neq 0 \qquad t = \frac{\hat\beta_4 - 0}{\hat\sigma_{\hat\beta_4}} \overset{H_0}{\sim} t(T-k)$$

Since $\mid t \mid = 2.717 > 1.99547 = t_{0.025}(68)$, the null hypothesis is rejected at the 5% significance level. Therefore, the variable h*aving free access to WiFi* is statistically significant.

c. Test the joint significance of the variables *proximity to a natural park*, *proximity to a beach*, *proximity to a lake or reservoir.*

$$H_0 : \beta_5 = \beta_6 = \beta_7 = 0 \qquad H_a : \exists \beta_i \neq 0 \qquad F = \frac{SSR_R - SSR_{UR}}{SSR_{UR}} \frac{T-k}{q} \overset{H_0}{\sim} \mathcal{F}(q, T-k)$$

Given that $F = 5.76258 > 2.7395 = \mathcal{F}_{0.05}(3, 68)$, the null hypothesis is rejected at the 5% significance level. Therefore, the variables *proximity to a natural park*, *proximity to a beach*, *proximity to a lake or reservoir* are jointly significant.

d. Test the individual significance of the variables *proximity to a natural park*, *proximity to a beach*, *proximity to a lake or reservoir.*

$$H_0 : \beta_i = 0 \qquad H_a : \beta_i \neq 0 \qquad t = \frac{\hat{\beta}_i - 0}{\hat{\sigma}_{\hat{\beta}_i}} \overset{H_0}{\sim} t(T-k)$$

$i = 5$: Since $\mid t \mid = 0.8476 < 1.99547 = t_{0.025}(68)$, the null hypothesis is not rejected at the 5% significance level. Therefore, the variable *proximity to a natural park* is not individually significant.

$i = 6$: Since $\mid t \mid = 3.613 > 1.99547 = t_{0.025}(68)$, the null hypothesis is rejected at the 5% significance level. Therefore, the variable *proximity to a beach* is individually significant.

$i = 7$: Since $\mid t \mid = 1.539 < 1.99547 = t_{0.025}(68)$, the null hypothesis is not rejected at the 5% significance level. Therefore, the variable *proximity to a lake or reservoir* is not individually significant.

e. Estimation results for the augmented model (including the variable *location*).

Model 4: OLS, using observations 1–75
Dependent variable: RP

	Coefficient	Std. Error	t-ratio	p-value
const	33.6327	6.80200	4.9445	0.0000
NR	1.31946	1.12334	1.1746	0.2443
BP	1.66487	1.08630	1.5326	0.1301
WIFIF	9.14238	3.35367	2.7261	0.0082
NPR	3.35482	3.96390	0.8463	0.4004
BER	16.3609	4.52546	3.6153	0.0006
LKR	11.6257	7.91546	1.4687	0.1466
LOCC	1.89981	4.55032	0.4175	0.6776

Mean dependent var	56.13893	S.D. dependent var	14.98446
Sum squared resid	11520.47	S.E. of regression	13.11287
R^2	0.306644	Adjusted R^2	0.234204
$F(7, 67)$	4.233071	P-value(F)	0.000641
Log-likelihood	−295.2101	Akaike criterion	606.4202
Schwarz criterion	624.9601	Hannan–Quinn	613.8230

Test the statistical significance of the variable *location.*

$$H_0 : \beta_8 = 0 \qquad H_a : \beta_8 \neq 0 \qquad t = \frac{\hat{\beta}_8 - 0}{\hat{\sigma}_{\hat{\beta}_8}} \overset{H_0}{\sim} t(T-k)$$

Since $\mid t \mid = 0.4175 < 1.99601 = t_{0.025}(67)$, the null hypothesis is not rejected at the 5% significance level. Therefore, the variable *location* is not statistically significant: the price of a room is the same if the holiday cottage is in the town center or far from it.

f. Given that the variable *location* is not statistically significant, the properties of the OLS estimator used in the previous model are not affected: it is linear, unbiased and it has the smallest variance in the class of linear and unbiased estimators.

g. The variables number of rooms, price of breakfast, *proximity to a natural park*, *proximity to a lake or reservoir* and *location* are not individually significant. Let's test whether they are jointly significant.

$$H_0 : \beta_2 = \beta_3 = \beta_5 = \beta_7 = \beta_8 = 0 \qquad F = \frac{SSR_R - SSR_{UR}}{SSR_{UR}} \frac{T-k}{q} \overset{H_0}{\sim} \mathcal{F}(q, T-k)$$
$$H_a : \exists \beta_i \neq 0$$

Since $F = 1.52656 < 2.35166 = \mathcal{F}_{0.05}(5, 67)$, the null hypothesis is not rejected at the 5% significance level. Therefore, the variables number of rooms, price of breakfast, *proximity to a natural park*, *proximity to a lake or reservoir* and *location* are not jointly significant.

The restricted model is:

$$RP_i = \beta_1 + \beta_4 \, WIFIF_i + \beta_6 \, BER_i + u_i$$

The estimation results are:

Model 5: OLS, using observations 1–75
Dependent variable: RP

	Coefficient	Std. Error	t-ratio	p-value
const	47.9589	2.43513	19.6946	0.0000
WIFIF	11.2175	3.15271	3.5580	0.0007
BER	16.0018	4.45227	3.5941	0.0006

Mean dependent var	56.13893	S.D. dependent var	14.98446
Sum squared resid	12832.91	S.E. of regression	13.35046
R^2	0.227656	Adjusted R^2	0.206202
$F(2, 72)$	10.61133	P-value(F)	0.000091
Log-likelihood	−299.2559	Akaike criterion	604.5118
Schwarz criterion	611.4642	Hannan–Quinn	607.2878

The estimator used is linear, unbiased (because the restrictions are true) and its variance is smaller than the variance of the estimator used in the unrestricted model.

Exercise E6.3 Soy milk

Part I. General Linear Regression Model.

a. Regression model: $S_t = \beta_1 + \beta_2 P_t + \beta_3 AE_t + \beta_4 AE_t^2 + u_t$

Model 2: OLS, using observations 1990:01–2012:06 ($T = 270$)
Dependent variable: S

	Coefficient	Std. Error	t-ratio	p-value
const	57.5860	43.7095	1.3175	0.1888
P	−1.38907	0.0807449	−17.2032	0.0000
AE	3.77879	0.610062	6.1941	0.0000
sq_AE	−0.0166627	0.00229460	−7.2617	0.0000

Mean dependent var	120.8475	S.D. dependent var	16.87912
Sum squared resid	28697.27	S.E. of regression	10.38674
R^2	0.625554	Adjusted R^2	0.621331
$F(3, 266)$	148.1277	P-value(F)	1.89e–56
Log-likelihood	−1013.042	Akaike criterion	2034.083
Schwarz criterion	2048.477	Hannan–Quinn	2039.863
$\hat{\rho}$	0.965596	Durbin–Watson	0.084094

SRF: $\hat{S}_t = 57.5860 - 1.38907 P_t + 3.77879\, AE_t - 0.0166627\, AE_t^2$

b. Overall significance test.

$$H_0 : \beta_2 = \beta_3 = \beta_4 = 0 \qquad H_a : \exists \beta_i \neq 0 \qquad F = \frac{R^2/(k-1)}{(1-R^2)/(T-k)} \overset{H_0}{\sim} \mathcal{F}(k-1, T-k)$$

Since $F = 148.1277 > 2.63854 = \mathcal{F}_{0.05}(3, 266)$, the null hypothesis is rejected at the 5% significance level. Therefore, the explanatory variables are jointly significant.

Test the individual significance of the variable price.

$$H_0 : \beta_2 = 0 \qquad H_a : \beta_2 \neq 0 \qquad t = \frac{\hat{\beta}_2 - 0}{\hat{\sigma}_{\hat{\beta}_2}} \overset{H_0}{\sim} t(T-k)$$

Since $\mid t \mid = 17.20 > 1.96892 = t_{0.025}(266)$, the null hypothesis is rejected at the 5% significance level. Therefore, the variable price is individually significant.

Test the individual significance of the variable advertising expenditures.

$$H_0 : \beta_3 = \beta_4 = 0 \qquad H_a : \exists \beta_i \neq 0 \qquad F = \frac{SSR_R - SSR_{UR}}{SSR_{UR}} \frac{T-k}{q} \overset{H_0}{\sim} \mathcal{F}(q, T-k)$$

Since $F = 92.5594 > 3.02973 = \mathcal{F}_{0.05}(2, 266)$, the null hypothesis is rejected at the 5% significance level. Therefore, the variable advertising expenditures is individually significant.

c. Two-sided test.

$$H_0 : \beta_4 = 0 \qquad t = \frac{\hat{\beta}_4 - 0}{\hat{\sigma}_{\hat{\beta}_4}} \overset{H_0}{\sim} t(T-k)$$
$$H_a : \beta_4 \neq 0$$

Since $\mid t \mid = 7.262 > 1.96892 = t_{0.025}(266)$, the null hypothesis is rejected at the 5% significance level. Therefore, the relationship between sales and advertising expenditures is not linear but quadratic.

d. Test a linear restriction.

$$H_0 : \beta_2 = \tfrac{-0.750}{50} = -0.015 \qquad F = \frac{SSR_R - SSR_{UR}}{SSR_{UR}} \frac{T-k}{q} \overset{H_0}{\sim} \mathcal{F}(q, T-k)$$
$$H_a : \beta_2 \neq -0.015$$

Since $F = 289.593 > 3.87666 = \mathcal{F}_{0.05}(1, 266)$, the null hypothesis is rejected at the 5% significance level. Therefore, a 50 cents increase in the price does not generate a decrease in sales of €750.

e. Prediction interval.

$CI(S_p)_{0.95} = [13.5156\,;\, 71.7971]$

If the price is 75 cents and the advertising expenditures are €20000, sales would lie between 13.5156 and 71.7971 thousands of euros. Therefore, a sale of 25 thousands of euros would be feasible.

Part II. Trend and seasonality.

a. Regression model:

$$S_t = \beta_1 + \beta_2\, P_t + \beta_3\, AE_t + \beta_4\, AE_t^2 + \beta_5\, time + \beta_6\, dm1t + \beta_7\, dm2_t + \ldots + \beta_{16}\, dm11_t + u_t$$

Model 4: OLS, using observations 1990:01–2012:06 (T = 270)
Dependent variable: S

	Coefficient	Std. Error	t-ratio	p-value
const	−384.992	23.5086	−16.3766	0.0000
P	2.65667	0.126825	20.9475	0.0000
AE	2.16316	0.274338	7.8850	0.0000
sq_AE	−0.00628330	0.00106251	−5.9136	0.0000
time	0.496684	0.0149453	33.2334	0.0000
dm1	−0.353509	1.36885	−0.2583	0.7964
dm2	−0.891238	1.36890	−0.6511	0.5156
dm3	−0.872756	1.36878	−0.6376	0.5243
dm4	−1.21630	1.36889	−0.8885	0.3751
dm5	−1.63315	1.36942	−1.1926	0.2341
dm6	−2.12044	1.36981	−1.5480	0.1229
dm7	−1.53919	1.38388	−1.1122	0.2671
dm8	−1.90023	1.38350	−1.3735	0.1708
dm9	−1.51783	1.38333	−1.0972	0.2736
dm10	−1.08713	1.38346	−0.7858	0.4327
dm11	−0.827367	1.38321	−0.5982	0.5503

Mean dependent var	120.8475	S.D. dependent var	16.87912
Sum squared resid	5345.186	S.E. of regression	4.587378
R^2	0.930255	Adjusted R^2	0.926137
$F(15, 254)$	225.8569	P-value(F)	2.6e–137
Log-likelihood	−786.1599	Akaike criterion	1604.320
Schwarz criterion	1661.895	Hannan–Quinn	1627.439
$\hat{\rho}$	0.815407	Durbin–Watson	0.367882

b. Estimated sales for December:

$$\hat{S}_{December} = -385.345509 + 2.65667\, P_t + 2.16316\, AE_t - 0.00628330\, AE_t^2 + 0.496684\, time$$

Estimated sales for August:

$$\hat{S}_{August} = -386.89223 + 2.65667\, P_t + 2.16316\, AE_t - 0.00628330\, AE_t^2 + 0.496684\, time - 1.90023$$

c. Test the statistical significance of the trend variable.

$$H_0 : \beta_5 = 0 \qquad t = \frac{\hat{\beta}_5 - 0}{\hat{\sigma}_{\hat{\beta}_5}} \overset{H_0}{\sim} t(T-k)$$
$$H_a : \beta_5 \neq 0$$

Since $\mid t \mid = 33.23 > 1.96935 = t_{0.025}(254)$, the null hypothesis is rejected at the 5% significance level. Therefore, the trend variable is statistically significant.

d. Test the statistical significance of seasonality.

$$H_0 : \beta_6 = \ldots \beta_{16} = 0 \qquad F = \frac{SSR_R - SSR_{UR}}{SSR_{UR}} \frac{T-k}{q} \overset{H_0}{\sim} \mathcal{F}(q, T-k)$$
$$H_a : \exists \beta_i \neq 0$$

Given that $F = 0.410257 < 1.82647 = \mathcal{F}_{0.05}(11, 254)$, the null hypothesis is not rejected at the 5% significance level. Therefore, seasonality is not a statistically significant variable.

e. Estimation results of the augmented model.

Model 5: OLS, using observations 1990:01–2012:06 ($T = 270$)
Dependent variable: S

	Coefficient	Std. Error	t-ratio	p-value
const	−250.520	30.6844	−8.1644	0.0000
P	−1.06176	0.605467	−1.7536	0.0807
AE	2.66103	0.267845	9.9350	0.0000
sq_AE	−0.00772453	0.00101699	−7.5954	0.0000
time	0.513666	0.0141954	36.1855	0.0000
dm1	−0.310565	1.27623	−0.2433	0.8079
dm2	−0.917931	1.27627	−0.7192	0.4727
dm3	−0.884234	1.27615	−0.6929	0.4890
dm4	−1.20676	1.27626	−0.9455	0.3453
dm5	−1.58079	1.27677	−1.2381	0.2168
dm6	−2.03810	1.27718	−1.5958	0.1118
dm7	−1.66481	1.29038	−1.2902	0.1982
dm8	−1.98003	1.28994	−1.5350	0.1260
dm9	−1.52685	1.28972	−1.1839	0.2376
dm10	−1.12275	1.28985	−0.8704	0.3849
dm11	−0.841019	1.28960	−0.6522	0.5149
sq_P	0.0191809	0.00306307	6.2620	0.0000

Mean dependent var	120.8475	S.D. dependent var	16.87912
Sum squared resid	4627.906	S.E. of regression	4.276929
R^2	0.939614	Adjusted R^2	0.935796
$F(16, 253)$	246.0466	P-value(F)	5.1e–144
Log-likelihood	−766.7075	Akaike criterion	1567.415
Schwarz criterion	1628.588	Hannan–Quinn	1591.980
$\hat{\rho}$	0.791304	Durbin–Watson	0.428458

Two-sided test:

$$\begin{array}{l} H_0 : \beta_{17} = 0 \\ H_a : \beta_{17} \neq 0 \end{array} \qquad t = \frac{\hat{\beta}_{17} - 0}{\hat{\sigma}_{\hat{\beta}_{17}}} \overset{H_0}{\sim} t(T-k)$$

Since $| t | = 6.2620 > 1.96938 = t_{0.025}(253)$, the null hypothesis is rejected at the 5% significance level. Therefore, the relationship between sales and prices is not linear but quadratic.

f. First, test whether seasonality is statistically significant in the augmented model. Since it is not statistically significant, specify the regression model:

$$S_t = \beta_1 + \beta_2\, P_t + \beta_3\, P_t^2 + \beta_4\, AE_t + \beta_5\, AE_t^2 + \beta_6\, time + u_t \quad t = 1990:1, \ldots, 2012:6$$

The estimation results are:

Model 6: OLS, using observations 1990:01–2012:06 ($T = 270$)
Dependent variable: S

	Coefficient	Std. Error	t-ratio	p-value
const	−251.137	30.3113	−8.2853	0.0000
P	−1.06533	0.598435	−1.7802	0.0762
sq_P	0.0191490	0.00302805	6.3239	0.0000
AE	2.66920	0.264313	10.0986	0.0000
sq_AE	−0.00776953	0.00100340	−7.7432	0.0000
time	0.512564	0.0140071	36.5931	0.0000

Mean dependent var	120.8475	S.D. dependent var	16.87912
Sum squared resid	4724.480	S.E. of regression	4.230338
R^2	0.938354	Adjusted R^2	0.937187
$F(5, 264)$	803.7083	P-value(F)	1.9e–157
Log-likelihood	−769.4957	Akaike criterion	1550.991
Schwarz criterion	1572.582	Hannan–Quinn	1559.661
$\hat{\rho}$	0.789293	Durbin–Watson	0.431438