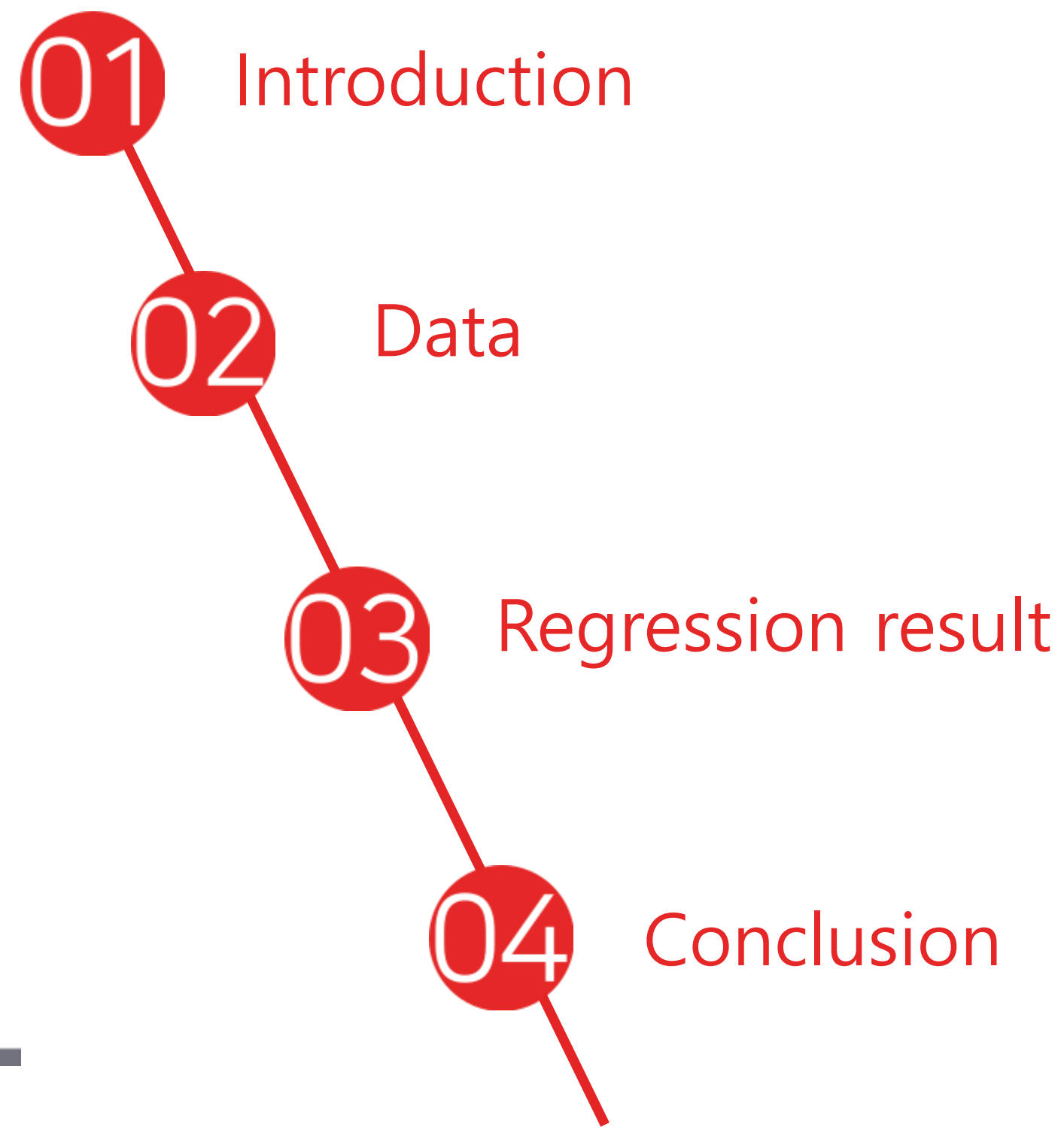
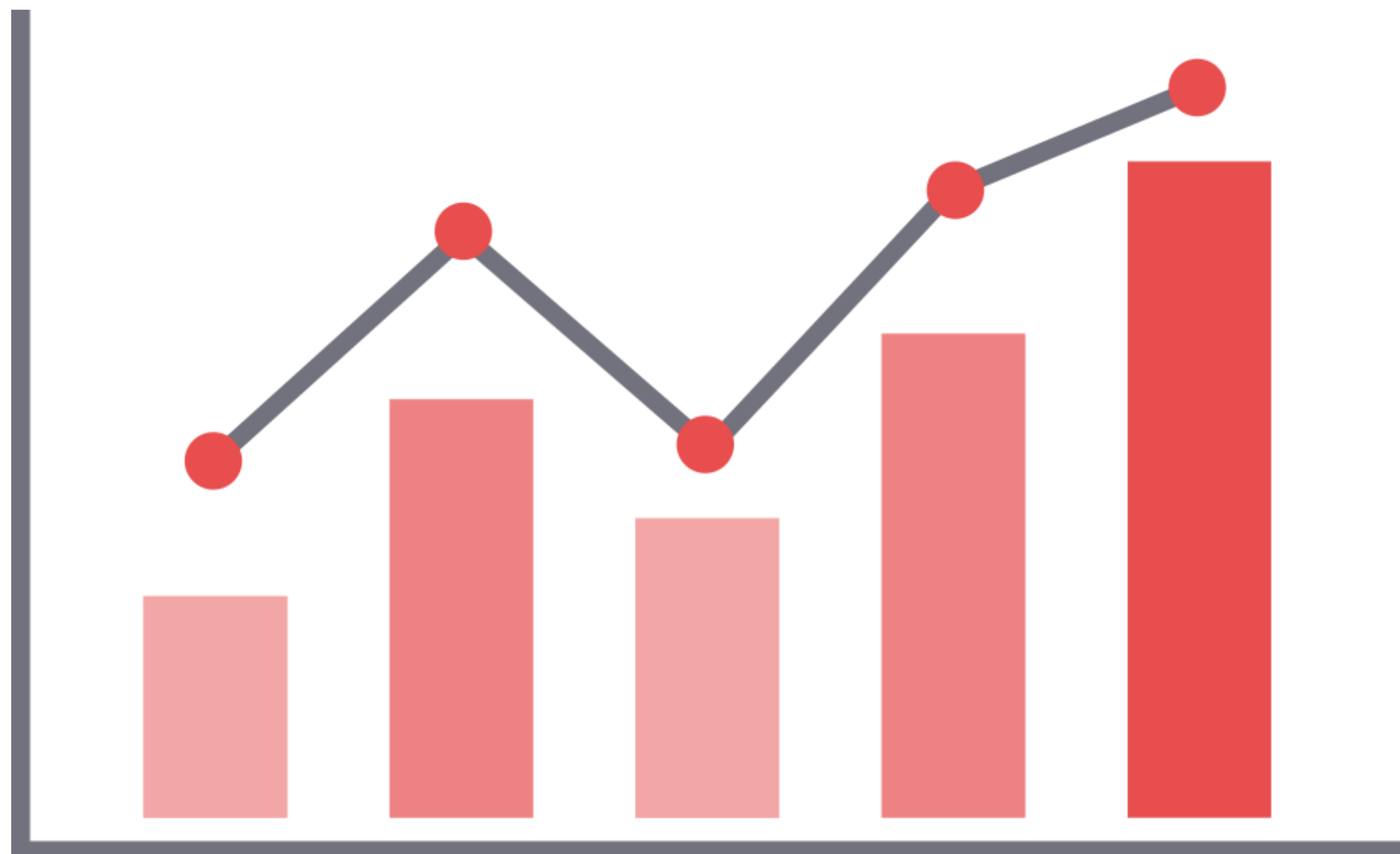


Analysis of the factors affecting the sales of venture companies

Team 7 Jiyeon Kim, Jihoo Kim, Myeongjin Woo



CONTENTS





01

Introduction

A grayscale background image of a city skyline with several tall skyscrapers. Overlaid on the image are several white, glowing arcs that connect different points, resembling a network or data flow diagram.

Reasons for selecting a topic

the factors affecting the sales of venture companies

Venture Company?

A grayscale background image of a city skyline with several tall skyscrapers. Overlaid on the image are white, glowing arcs that connect different points, resembling a network or data flow diagram.

Review of previous studies

Previous studies said the most important factor for Venture companies is ...

Entrepreneurial Spirit



02

Data

Introduction of Data

Detailed Survey of Venture Businesses(2020)

구분_법인/가중치		[쿼터] 지역	[쿼터] 업종	[쿼터] 규모	A1-4. 벤처	A1-5. 회사	A1-6. KSIC	[쿼터] 업력	A2-2. 기업	A2-3. 제조	A2-3. 일부	B2-1. 창업	B2-2. 창업	B2-3. 창업	B2-3-1. 창	B2-4. 창업
개인	0.079104	1	8	3	4	경기	62021	3	3	4		1	1	3	3	10
개인	0.079104	3	8	3	2	경남	62021	3	4	1		1	2	4	1	6
개인	0.431785	1	10	3	1	경기	1299	2	3	1		1	2	5		6
개인	0.431785	3	10	3	2	경남	42499	3	4	4		1	2	3	1	6
개인	0.450524	2	10	2	2	강원	42412	3	4	1		1	2	3	1	6
개인	0.450524	4	10	2	2	대구	42201	4	4	4		1	2	3	1	5
개인	0.512647	1	5	6	2	경기	29192	4	3	2	30	1	3	3	1	6
개인	0.512647	1	5	6	2	인천	29299	4	4	2	10	1	3	3	1	6
개인	0.512647	3	5	6	2	부산	29120	3	4	2	10	1	3	4	1	6
개인	0.512647	3	5	6	2	부산	29132	4	4	2	25	1	2	5		6
개인	0.512647	4	5	6	2	경북	30332	4	5	1		1	1	3	1	6
개인	0.539731	1	10	3	1	서울	58113	4	4	2	25	1	2	3	4	8
개인	0.539731	1	10	3	2	서울	58122	4	4	1		2	1	2	1	1
개인	0.539731	2	10	3	2	대전	42311	4	4	4		1	2	1	1	6
개인	0.539731	4	10	3	2	경북	1231	2	4	4		1	2	3	4	6
개인	0.63518	2	10	4	1	대전	1299	2	3	1		1	2	3	3	9
개인	0.672358	1	9	4	2	경기	18119	4	4	1		1	2	5		6
개인	0.672358	4	9	4	1	대구	18119	3	4	1		1	1	3	1	6
개인	0.711906	1	9	5	1	경기	46315	2	3	2	20	1	2	4	5	6
개인	0.711906	1	9	5	1	경기	46413	2	2	3		2	2	5		6
개인	0.711906	1	9	5	2	인천	18119	4	5	2	30	1	3	5		7
개인	0.711906	2	9	5	1	강원	46322	2	3	2	50	2	4	5		6

Variables

X Variables(Independent variables)

- ✓ 1. Number of College graduate employees (대졸인력)
- ✓ 2. Business Founder's year of practical Experience (실무년수)
- ✓ 3. number of permanent employee (정규직)
- ✓ 4. The stage of corporate Growth (기업성장단계)
 - ✓ 5. Market Share (시장점유율)
 - ✓ 6. Amount of Liabilities (부채)
 - ✓ 7. Structure of quota (지분 구조)
- ✓ 8. Number of industrial property right (산업재산권)

Y Variable
(Dependent
variables)

Sales Scale

Data preprocessing

Final Data after preprocessing

	규모	기업성장단계	실무년수	정규직	대졸인력	지분구조	부채	산업재산권	시장점유율
count	2500.000000	2500.000000	2500.000000	2500.000000	2500.000000	2500.000000	2500.000000	2500.000000	2500.000000
mean	114.850400	3.261200	10.531960	47.598800	21.310800	56.468720	8755.375200	11.918800	4.770828
std	133.526296	0.824892	6.982106	64.395016	32.537438	29.362204	18476.688633	23.439709	9.256763
min	1.000000	1.000000	0.000000	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	20.000000	3.000000	5.000000	10.000000	4.000000	40.000000	933.000000	1.000000	0.100000
50%	80.000000	3.000000	10.000000	23.000000	10.000000	55.000000	3008.500000	4.000000	1.000000
75%	120.000000	4.000000	15.000000	57.000000	24.000000	80.000000	7883.000000	12.000000	5.000000
max	500.000000	5.000000	40.000000	658.000000	490.000000	100.000000	380473.000000	230.000000	80.000000

03

Regression Result



1. Simple Linear Regression

Variable Selection:
number of permanent employee

Hypothesis Testing with
Heteroskedastic
Robust Standard Error:
P-value = 0.000 < 0.05

An increase of 1 unit of the variable
makes an increase of 1.4042 unit
of sales.

$R^2 = 0.459$

OLS Regression Results

=====						
Dep. Variable:	규모		R-squared:	0.459		
Model:	OLS		Adj. R-squared:	0.458		
Method:	Least Squares		F-statistic:	438.4		
Date:	Tue, 07 Jun 2022		Prob (F-statistic):	8.18e-90		
Time:	04:10:05		Log-Likelihood:	-15016.		
No. Observations:	2500		AIC:	3.004e+04		
Df Residuals:	2498		BIC:	3.005e+04		
Df Model:	1					
Covariance Type:	HC0					
=====						
	coef	std err	z	P> z	[0.025	0.975]

const	48.0107	2.878	16.683	0.000	42.370	53.651
정규직	1.4042	0.067	20.937	0.000	1.273	1.536
=====						
Omnibus:	919.297		Durbin-Watson:	0.808		
Prob(Omnibus):	0.000		Jarque-Bera (JB):	5344.531		
Skew:	1.629		Prob(JB):	0.00		
Kurtosis:	9.379		Cond. No.	99.6		
=====						

Notes:

[1] Standard Errors are heteroscedasticity robust (HC0)

2. Multiple Linear Regression

Method: Backward Elimination

At first, all variables are selected.
 $Y \sim X1+X2+X3+X4+X5+X6+X7+X8$

Hypothesis Testing with Heteroskedastic
Robust Standard Error:
P-value of the variable
“Number of College graduate employees” > 0.05
-> eliminate the variable

Adjusted $R^2 = 0.525$

```
=====
                        OLS Regression Results
=====
Dep. Variable:          규모      R-squared:          0.527
Model:                  OLS      Adj. R-squared:       0.525
Method:                 Least Squares      F-statistic:       115.1
Date:                   Mon, 06 Jun 2022    Prob (F-statistic):  4.92e-164
Time:                   13:04:39           Log-Likelihood:     -14847.
No. Observations:       2500             AIC:              2.971e+04
Df Residuals:           2491             BIC:              2.976e+04
Df Model:               8
Covariance Type:        HC0
=====
                        coef      std err          z      P>|z|      [0.025      0.975]
-----
Intercept              34.9885        9.298        3.763      0.000       16.765       53.212
실무년수              -0.4284         0.233       -1.838      0.066       -0.885        0.028
정규직                0.9026         0.138        6.556      0.000         0.633        1.172
대졸인력              0.0736         0.200         0.368      0.713       -0.318        0.466
산업재산권            0.3100         0.138         2.247      0.025         0.040        0.580
기업성장단계          8.4997         2.233         3.806      0.000         4.123       12.876
부채                  0.0020         0.000         4.029      0.000         0.001        0.003
지분구조             -0.2317         0.075       -3.081      0.002       -0.379       -0.084
시장점유율            0.8455         0.253         3.346      0.001         0.350        1.341
=====
Omnibus:                853.675    Durbin-Watson:       0.903
Prob(Omnibus):          0.000    Jarque-Bera (JB):    5683.208
Skew:                   1.450    Prob(JB):            0.00
Kurtosis:               9.793    Cond. No.            1.10e+05
=====
```

Warnings:

```
[1] Standard Errors are heteroscedasticity robust (HC0)
[2] The condition number is large, 1.1e+05. This might indicate that there are
strong multicollinearity or other numerical problems.
```


2. Multiple Linear Regression

Method: Backward Elimination

Eliminated Variable:

1. Number of College graduate employees

Hypothesis Testing with Heteroskedastic

Robust Standard Error:

P-value of the variable "Business Founder's year of practical experience" > 0.05

-> eliminate the variable

Adjusted $R^2 = 0.525$

(equal to the previous model)

```
=====
                        OLS Regression Results
=====
Dep. Variable:          규모      R-squared:          0.527
Model:                  OLS      Adj. R-squared:       0.525
Method:                 Least Squares      F-statistic:       131.9
Date:                  Mon, 06 Jun 2022      Prob (F-statistic): 1.62e-165
Time:                  12:18:58      Log-Likelihood:    -14847.
No. Observations:      2500      AIC:              2.971e+04
Df Residuals:          2492      BIC:              2.976e+04
Df Model:              7
Covariance Type:       HC0
=====

               coef      std err          z      P>|z|      [0.025      0.975]
-----
Intercept      35.6815      9.056      3.940      0.000      17.933      53.430
실무년수      -0.4426      0.236     -1.880      0.060     -0.904      0.019
정규직         0.9366      0.095      9.893      0.000      0.751      1.122
산업재산권      0.3045      0.137      2.229      0.026      0.037      0.572
기업성장단계      8.4239      2.195      3.837      0.000      4.121     12.726
부채           0.0020      0.000      4.000      0.000      0.001      0.003
지분구조      -0.2334      0.075     -3.097      0.002     -0.381     -0.086
시장점유율      0.8264      0.242      3.416      0.001      0.352      1.301
=====

Omnibus:          856.998      Durbin-Watson:       0.902
Prob(Omnibus):    0.000      Jarque-Bera (JB):    5751.491
Skew:            1.454      Prob(JB):            0.00
Kurtosis:        9.838      Cond. No.            1.10e+05
=====
```

Warnings:

[1] Standard Errors are heteroscedasticity robust (HC0)

[2] The condition number is large, 1.1e+05. This might indicate that there are strong multicollinearity or other numerical problems.

2. Multiple Linear Regression

Final Model of Multiple Linear Regression!

Eliminated Variable:

1. Number of College graduate employees
2. Business Founder's year of practical experience

Hypothesis Testing with
Heteroskedastic

Robust Standard Error:

P-value of all remained variables < 0.05

Adjusted $R^2 = 0.525$
(equal to the previous model)

```
=====
                        OLS Regression Results
=====
Dep. Variable:          규모      R-squared:                0.526
Model:                  OLS      Adj. R-squared:           0.525
Method:                 Least Squares      F-statistic:            152.4
Date:                  Mon, 06 Jun 2022    Prob (F-statistic):      4.25e-165
Time:                  13:05:07           Log-Likelihood:         -14849.
No. Observations:      2500              AIC:                   2.971e+04
Df Residuals:          2493              BIC:                   2.975e+04
Df Model:               6
Covariance Type:       HC0
=====
                        coef      std err          z      P>|z|      [0.025      0.975]
-----
Intercept             29.9727      8.480        3.535      0.000      13.353      46.592
정규직                 0.9399      0.095        9.905      0.000       0.754       1.126
산업재산권             0.3085      0.136        2.262      0.024       0.041       0.576
기업성장단계           8.6713      2.205        3.933      0.000       4.350      12.993
부채                   0.0020      0.000        3.988      0.000       0.001       0.003
지분구조              -0.2305      0.075       -3.057      0.002      -0.378      -0.083
시장점유율             0.8055      0.239        3.365      0.001       0.336       1.275
=====
Omnibus:               860.983      Durbin-Watson:          0.902
Prob(Omnibus):         0.000      Jarque-Bera (JB):       5809.615
Skew:                  1.460      Prob(JB):               0.00
Kurtosis:              9.874      Cond. No.                1.03e+05
=====
```

Warnings:

- [1] Standard Errors are heteroscedasticity robust (HC0)
- [2] The condition number is large, 1.03e+05. This might indicate that there are strong multicollinearity or other numerical problems.



3. Nonlinear Regression : Interaction

2. Interaction of (The stage of corporate Growth)*(Market Share)

P-value of Interaction of (The stage of corporate Growth)*(Market Share) > 0.05
-> eliminate the Interaction

Adjusted $R^2 = 0.543$
(little bit larger than multiple linear regression)

Dep. Variable:		규모	R-squared:	0.545		
Model:		OLS	Adj. R-squared:	0.543		
Method:		Least Squares	F-statistic:	143.7		
Date:		Mon, 06 Jun 2022	Prob (F-statistic):	7.62e-238		
Time:		15:07:41	Log-Likelihood:	-14798.		
No. Observations:		2500	AIC:	2.962e+04		
Df Residuals:		2489	BIC:	2.968e+04		
Df Model:		10				
Covariance Type:		HC0				
=====						
	coef	std err	z	P> z	[0.025	0.975]

Intercept	23.8481	9.207	2.590	0.010	5.803	41.893
실무년수	-0.1946	0.231	-0.841	0.400	-0.648	0.259
정규직	1.0736	0.134	8.011	0.000	0.811	1.336
대졸인력	0.8491	0.249	3.413	0.001	0.361	1.337
산업재산권	0.2257	0.130	1.741	0.082	-0.028	0.480
기업성장단계	5.7796	2.222	2.601	0.009	1.425	10.135
부채	0.0018	0.000	3.974	0.000	0.001	0.003
지분구조	-0.1418	0.074	-1.926	0.054	-0.286	0.002
시장점유율	0.3908	0.603	0.648	0.517	-0.792	1.573
정규직:대졸인력	-0.0036	0.001	-2.977	0.003	-0.006	-0.001
기업성장단계:시장점유율	0.1415	0.197	0.719	0.472	-0.244	0.527
=====						
Omnibus:	876.738		Durbin-Watson:	0.852		
Prob(Omnibus):	0.000		Jarque-Bera (JB):	4973.053		
Skew:	1.551		Prob(JB):	0.00		
Kurtosis:	9.174		Cond. No.	1.22e+05		
=====						

```
[1] Standard Errors are heteroscedasticity robust (HC0)
[2] The condition number is large, 1.22e+05. This might indicate that there are
strong multicollinearity or other numerical problems.
```


3. Nonlinear Regression : Interaction

Method: Backward Elimination

Eliminated Variable:

1. Interaction of (The stage of corporate Growth)*(Market Share)

P-value of “Business Founder’s year of practical Experience” & “Number of industrial property right”
> 0.05
-> eliminate two variables

Adjusted R² = 0.543
(equal to the previous model)

OLS Regression Results						
=====						
Dep. Variable:	규모	R-squared:	0.545			
Model:	OLS	Adj. R-squared:	0.543			
Method:	Least Squares	F-statistic:	159.0			
Date:	Mon, 06 Jun 2022	Prob (F-statistic):	5.85e-238			
Time:	15:07:44	Log-Likelihood:	-14798.			
No. Observations:	2500	AIC:	2.962e+04			
Df Residuals:	2490	BIC:	2.967e+04			
Df Model:	9					
Covariance Type:	HC0					
=====						
	coef	std err	z	P> z	[0.025	0.975]

Intercept	21.6419	9.063	2.388	0.017	3.879	39.405
실무년수	-0.1956	0.231	-0.845	0.398	-0.649	0.258
정규직	1.0750	0.134	8.046	0.000	0.813	1.337
대졸인력	0.8456	0.248	3.404	0.001	0.359	1.332
산업재산권	0.2282	0.129	1.764	0.078	-0.025	0.482
기업성장단계	6.4284	2.149	2.991	0.003	2.216	10.641
부채	0.0018	0.000	3.982	0.000	0.001	0.003
지분구조	-0.1410	0.074	-1.914	0.056	-0.285	0.003
시장점유율	0.8564	0.239	3.588	0.000	0.389	1.324
정규직:대졸인력	-0.0035	0.001	-2.980	0.003	-0.006	-0.001
=====						
Omnibus:	876.695	Durbin-Watson:	0.853			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	4957.825			
Skew:	1.552	Prob(JB):	0.00			
Kurtosis:	9.161	Cond. No.	1.15e+05			
=====						

Warnings:

- [1] Standard Errors are heteroscedasticity robust (HC0)
- [2] The condition number is large, 1.15e+05. This might indicate that there are strong multicollinearity or other numerical problems.

3. Nonlinear Regression : Interaction

Final Model of Nonlinear Regression!

- Eliminated Variable:
- 1. Interaction of (The s Growth)*(Market Sha
 - 2. Business Founder's Experience
 - 3. Number of industrial

P-value of all remained
-> Stop eliminating and

Adjusted R^2 = 0.542

Sales scale = (Number of permanent employees) + (Number of College graduate Employees) + (The stage of corporate Growth) + (Amount of Liabilities) + (Structure of quota) + (Market Share) + Interaction of (Number of permanent employees)*(Number of College graduate Employees)

OLS Regression Results			
=====			
Dep. Variable:	규모	R-squared:	0.544
Model:	OLS	Adj. R-squared:	0.542
		F-statistic:	199.2
		Prob (F-statistic):	4.66e-235
		Log-Likelihood:	-14802.
			2.962e+04
			2.967e+04
=====			
	P> z	[0.025	0.975]

	0.024	2.537	35.698
2	0.000	0.850	1.367
7	0.001	0.358	1.348
42	0.001	2.772	11.247
	0.000	0.001	0.003
0	0.038	-0.295	-0.009
3	0.000	0.425	1.359
34	0.002	-0.006	-0.001
=====			
	in-Watson:		0.848
	ue-Bera (JB):		4882.049
	(JB):		0.00
	. No.		1.07e+05
=====			

Warnings:

[1] Standard Errors are heteroscedasticity robust (HC0)

[2] The condition number is large, 1.07e+05. This might indicate that there are strong multicollinearity or other numerical problems.



04

Conclusion



Conclusion

Linear Model Analyze

- 1. Coefficient
- 2. P-value
- 3. Confidence Interval
[0.025, 0.975]

Model 4 has the best explanatory power:
BIC: 2.975e+04
Adjusted R^2 = 0.525

Regressor	(1)	(2)	(3)	(4)
정규직	1.4042 (0.000) [1.273,1.536]	0.9026 (0.000)	0.9366 (0.000) [0.751,1.122]	0.9399 (0.000) [0.754,1.126]
실무년수		-0.4284 (0.066) [0.885,0.028]	-0.4426 (0.060) [-0.904,0.019]	
대졸인력		0.0736 (0.713) [0.318,0.466]		
산업재산권		0.3100 (0.025) [0.040,0.580]	0.3045 (0.026) [0.037,0.572]	0.3085 (0.024) [0.041,0.576]
기업성장단계		8.4997 (0.000) [4.123,12.876]	8.4239 (0.000) [4.121,12.726]	8.6713 (0.000) [4.350,12.993]
부채		0.0020 (0.000) [0.001,0.003]	0.0020 (0.000) [0.001,0.003]	0.0020 (0.000) [0.001,0.003]
지분구조		-0.2317 (0.002) [-0.379,0.084]	-0.2334 (0.002) [-0.381,0.086]	-0.2305 (0.002) [-0.378,0.083]
시장점유율		0.8455 (0.001) [0.350,1.341]	0.8264 (0.001) [0.352,1.301]	0.8055 (0.001) [0.336,1.275]
Intercept	48.0107	34.9885	35.6815	29.9727
Summary Statistics				
BIC	3.005e+04	2.976e+04	2.976e+04	2.975e+04
Adj. R-squared	0.458	0.525	0.525	0.525
n	2500	2500	2500	2500



Non-Linear Model Analyze

- 1. Coefficient
- 2. P-value
- 3. Confidence Interval
[0.025, 0.975]

Model 8 has the best explanatory power:
BIC: 2.967e+04
Adjusted R^2 = 0.542

Regressor	(5)	(6)	(7)	(8)
정규직	1.0736	1.0750	1.0748	1.1087
	(0.000)	(0.000)	(0.000)	(0.000)
	[0.811,1.336]	[0.813,1.337]	[0.813,1.337]	[0.850,1.367]
실무년수	-0.1946	-0.1956		
	(0.400)	(0.398)		
	[-0.648,0.259]	[-0.649,0.258]		
대출인력	0.8491	0.8456	0.8566	0.8530
	(0.001)	(0.001)	(0.001)	(0.001)
	[0.361,1.337]	[0.359 ,1.332]	[0.371,1.342]	[0.358,1.348]
산업재산권	0.2257	0.2282	0.2298	
	(0.082)	(0.078)	(0.075)	
	[-0.028,0.480]	[-0.025,0.482]	[-0.023,0.483]	
기업성장단계	5.7796	6.4284	6.5283	7.0093
	(0.009)	(0.003)	(0.002)	(0.001)
	[1.425,10.135]	[2.216,10.641]	[2.322,10.735]	[2.772,11.247]
부채	0.0018	0.0018	0.0018	0.0019
	(0.001)	(0.000)	(0.000)	(0.000)
	[0.001,0.003]	[0.001,0.003]	[0.001,0.003]	[0.001,0.003]
지분구조	-0.1418	-0.1410	-0.1390	-0.1520
	(0.054)	(0.056)	(0.059)	(0.038)
	[-0.286,0.002]	[-0.285,0.003]	[-0.283,0.005]	[-0.295,-0.009]
시장점유율	0.3908	0.8564	0.8489	0.8920
	(0.517)	(0.000)	(0.000)	(0.000)
	[-0.792,1.573]	[0.389,1.324]	[0.384,1.314]	[0.425,1.359]
정규직: 대출인력	-0.0036	-0.0035	-0.0036	-0.0037
	(0.003)	(0.003)	(0.003)	(0.002)
	[-0.006,-0.001]	[-0.006,-0.001]	[-0.006,-0.001]	[-0.006,-0.001]
기업성장단계: 시장점유율	0.1415			
	(0.472)			
	[0.244,0.527]			
Intercept	23.8481	21.6419	19.0100	19.1178
Summary Statistics				
BIC	2.968e+04	2.967e+04	2.967e+04	2.967e+04
Adj. R-squared	0.543	23.8481	0.543	0.542
n	2500	2500	2500	2500

Final Model

Sales scale = (Number of permanent employees) + (Number of College graduate Employees) + (The stage of corporate Growth) + (Amount of Liabilities) + (Structure of quota) + (Market Share) + Interaction of (Number of permanent employees)*(Number of College graduate Employees)

Positive effect

- ✓ number of permanent employee (정규직)
- ✓ 4. The stage of corporate Growth (기업성장단계)
 - ✓ 5. Market Share (시장점유율)
 - ✓ 6. Amount of Liabilities (부채)

Negative effect

- ✓ Structure of quota (지분구조)

Conclusion

The important factor for the success of a venture company is ...
not its individual ability, but its activities and consequent performance.



Individual Experience

The diagram consists of two ovals. The left oval is light red and contains the text 'Individual Experience'. The right oval is light blue and contains the text 'Structure and Activities of Bussiness'. A large, purple, outlined chevron arrow points from the red oval to the blue oval, indicating a relationship or flow from individual experience to business structure and activities.

Structure and Activities
of Bussiness



THANK YOU

감사합니다