

逢甲大學學生報告 ePaper

報告題名：

點燃你的肺-肺癌成因之探討

Burning your lungs—

The examination of the causes of lung cancer

作者：楊侑錚、范芸翎、王雨柔、張梓淇、陳羽均

系級：統計二甲

學號：D0680595、D0680565、D0609436、D0680429、D0636314

開課老師：劉峰旗 助理教授

課程名稱：統計計算

開課系所：統計學系

開課學年： 107 學年度 第一 學期

## 中文摘要

近幾年科技發展快速，各個領域隨之進步，醫療技術也不例外。雖然許多疾病已經可以被完全治癒，不過有些疾病卻還是不能完全被治癒，像是「癌症」。2018 年所有癌症中有著最高發病率以及致死率的癌症即為「肺癌」，所以引發本文探討我國肺癌成因的主要研究動機。本研究將使用我國 97 年到 106 年肺癌死亡人數資料分別與可能引起肺癌的因子進行迴歸分析及獨立樣本 t 檢定之探討，以了解肺癌的死亡與工廠數、全國 18 歲以上吸菸比率有無顯著的線性迴歸關係，並比較在空氣品質好與不好的情況下，肺癌發生率之差異情況，藉由此分析結果觀察什麼因素可能會影響肺癌，並可預先構想解決對策，以防止其發生。

關鍵字：肺癌、吸菸、空氣品質、迴歸分析、獨立樣本 t 檢定。

## Abstract

In recent years, science and technology have developed rapidly, and various fields have also improved. Medical technology is no exception. Although many diseases can be cured by advanced medical technology, some diseases cannot be completely cured, such as "cancer". The cancer with the highest incidence and mortality in all cancers in 2018 is "lung cancer", so it motivates us to explore the causes of the lung cancer in Taiwan. The regression analyses and an independent sample t-test are used for investigating the relationship of the number of deaths of lung cancer from 2008 to 2017, the number of factories, and the rate of smoking people above eighteen-year-old. The results of our analyses present that there is no significant linear relationship between the smoking rate and the number of deaths of lung cancer, but there are a significant linear relationship between the number of factories and the number of deaths of lung cancer, and a significant difference of the ratios of death between poor and good air qualities. Therefore, what factors may affect lung cancer and how to prevent the occurrence of lung cancer are investigated in our study.

**Key Words:** lung cancer, regression analysis, independent sample t-test.

## 目錄

摘要 .....	1
ABSTRACT.....	2
第一章、序論 .....	4
第一節、研究動機.....	4
第二節、研究目的.....	4
第二章、研究背景及肺癌基本介紹 .....	5
第一節、研究背景.....	5
第二節、肺癌基本介紹.....	5
第三章、研究分析方法 .....	7
第一節、肺癌死亡人數與工廠數的關係.....	8
第二節、肺癌死亡人數與 18 歲以上吸菸率的關係.....	8
第三節、肺癌死亡率與空氣汙染的關聯.....	8
第四章、結果分析與應用價值 .....	9
第一節、各縣市每年肺癌人數與各縣市每年工廠數的關聯(迴歸分析).....	9
第二節、全國每年肺癌人數與全國 18 歲以上吸菸比率的關聯(迴歸分析).....	11
第三節、每年肺癌人數與空氣汙染的關聯(雙樣本平均數差異 T 檢定).....	13
第五章、結論與討論 .....	13
第一節、結論.....	13
第二節、討論.....	14
第六章、團隊合作之學習與成長 .....	14
第七章、參考文獻與資料來源 .....	15

## 第一章、序論

### 第一節、研究動機

從小，我們就被教導導致罹患肺癌的原因有很多種，例如抽菸、吸二手菸、職業等因素。然而，隨著年紀增長，與以前不同的是，我們逐漸可以從各種資訊媒體中發現，例如遺傳、接觸油煙的時間、吸入二手菸、空氣污染和許多原因也都被提及，且印象最深刻的原因不外乎就是抽菸及吸二手菸。不過，令我們驚訝的是，曾經有醫學實驗上的數據顯示，不抽菸也會罹患肺癌，翻轉了我們以前在健康課本上及考試時對於抽菸造成肺癌的認知。因此，引發我們對於這個主題的研究動機，想要透過統計分析方法來瞭解哪些因素可能會造成肺癌發病率的提升。我們選擇 3 個較有疑問且數據較完整的變數，如各縣市工廠數、全國 18 歲以上吸菸比率、各縣市空氣品質，來驗證這些變數是否真的與罹患肺癌有關，以及對肺癌發生率是否有影響。

### 第二節、研究目的

隨著醫療的進步，許多疾病都可以被治癒，人們的平均壽命也因為發達的醫療而逐漸延長，不過還是有一些疾病仍然無法被完全醫治，像是「癌症」。根據我國衛福部以及各報導的顯示，「癌症」已經連續 35 年蟬聯我國 10 大死因之首，其中又以肺癌、乳癌、大腸直腸癌、攝護腺癌與胃癌為全球最為常見的五大癌症，且在民國 103 年十大癌症死因排名，肺癌位居第一。因為近幾年癌症人數因為某些因素大幅攀升，於是我們以肺癌作為主題，去探討影響肺癌發生率的成因，了解有什麼原因可能造成肺癌人數的上升，以下我們採用三項變數來進行我們的研究：各縣市每年工廠數，全國 18 歲以上吸菸比率與空氣品質，去觀察各變數與肺癌之間的關係。

## 第二章、研究背景及肺癌基本介紹

### 第一節、研究背景

由於從小的教育以及報章雜誌的資訊，所以一開始我們就先預期吸菸比率、各縣市工廠數以及空氣汙染的天數，皆有可能會造成每年肺癌人數上升，並進行資料的分析驗證。當然近幾年的資訊告訴我們，抽菸被認為是最大發生肺癌的危險因子，空氣汙染中的細懸浮微粒(PM2.5)可以穿越細支氣管壁直達肺泡，造成慢性發炎甚至引發癌症，工廠所排放廢氣中的鎘、砷、放射性氫氣、鎳化合物、鉻化合物、氯乙烯、柴油燃燒物、焦油化合物等，也都可能導致肺癌，這與我們傳統觀念及預期的結果是符合的，再來，我們將進行資料分析，驗證是否與我們預期的結果是一樣的。

### 第二節、肺癌基本介紹

肺癌是發生在支氣管上皮或肺泡上皮的癌症，因此也可以稱為原發性支氣管肺癌，肺癌大部分會先發生於支氣管接著再向肺泡發展，也有少數會先發生於肺泡再向支氣管蔓延。近幾年來，空氣汙染與吸菸人口的增加，成為罹患肺癌人數不停上升的重要因素。吸菸是引起肺癌的主要原因之一，男性吸菸者，得肺癌的機會為不吸菸者的十倍，而女性則是五倍，可能因為女性吸菸量較小的原因。只要停止吸菸，或許危害的機會就會減少。並且根據資料(臺北市立聯合醫院癌症防治研究發展中心癌症資訊網)顯示，有 25%不抽菸者得肺癌則是因為吸了二手菸所造成的，長時間接觸二手菸得肺癌的機會大概是不吸菸者的兩倍。除此之外，如礦苗的放射線、石棉、熱焦油煙氣、工廠煙塵、汙染空氣中的氧化碘、鎳、鉻化物、汽機車所排放廢氣中的苯及多環芳香烴等，也是導致肺癌的原因之一。華人女性肺腺癌比西方國家來的多，有人推斷可能與烹飪食物過程中產生之油煙以及燒香等習俗有關。以國泰醫院為例，2005 年肺癌病患中，男性病患與女性病患比例分別約為 70%與 30%，年齡層大部分集中於 60 歲至 79 歲之間，腫瘤期別多集中於第三期及第四期。

根據新光醫療財團法人-新光吳火獅紀念醫院胸腔外科顯示，肺癌又因為發生癌的上皮細胞不同，將其分為小細胞癌與非小細胞癌，小細胞的惡性度高，在其體積很小時便會轉移出去，所以只要確診後，便會採取化學治療為主。而後者會依細胞型態區分為：鱗狀細胞癌、腺癌(包括細支氣管肺泡癌也簡稱肺泡癌)、大細胞癌、腺鱗狀細胞癌、類癌、支氣管腺癌(包括腺樣囊性癌、黏液上皮樣癌)等，而這種細胞癌的治療以手術切除為主，放射及化學治療為輔。

表一、2018 年全球前 10 大癌症的新增病例數和死亡人數

癌症	新增病例 (所占百分比)	死亡人數 (所占百分比)
肺癌	2093876(11.6%)	1761007(18.4%)
乳腺癌	2088849(11.6%)	626679 (6.6%)
前列腺癌	1276106 (7.1%)	358989 (3.8%)
結直腸癌	1096601 (6.1%)	551269 (5.8%)
非黑色素瘤的皮膚癌	1042056 (5.8%)	65155 (0.7%)
胃癌	1033701 (5.7%)	782685 (8.2%)
肝癌	841080 (4.7%)	781631 (8.2%)
直腸癌	704376 (3.9%)	310394 (3.2%)
食道癌	572034 (3.2%)	508585 (5.3%)
子宮頸癌	569847 (3.2%)	311365 (3.3%)

資料來源：Bray *et al.* (2018) Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries, *CA CANCER J CLIN*, 68,394-424.

上方表一是由 Bray *et al.* (2018)發表之最新全球腫瘤統計分析結果，研究結果顯示，36 種癌症中發病率前四名為：第一名分別是肺癌以及女性乳腺癌，皆占癌症總發病人口的 11.6%，第三名則是前列腺癌，占癌症總發病人口的 7.1%，最後第四名為結直腸癌，占癌症總發病人口的 5.7%。而致死率前三名排行為：第一名仍然是肺癌，占癌症總發病人口的 18.4%，第二名是結直腸癌，占癌症總發病人口的 9.2%，第三名胃癌和肝癌的致死率，皆占癌症總發病人口的 8.2%。

### 第三章、研究分析方法

#### 第一節、資料蒐集

首先，我們由衛生福利部統計處提供的死因統計資料中得到各縣市肺癌死亡人數資料，這組資料共有 220 筆樣本，並且以縣市做為分類共可分成 22 組，而每組共有 10 年的資料。資料如下表二所示。

表二、2008 年至 2017 年的各縣市肺癌死亡人數

西元年	新北市	台北市	桃園市	台中市	台南市	高雄市	宜蘭縣	新竹縣	苗栗縣	彰化縣	南投縣
2008	889	950	498	634	986	1189	228	196	184	512	268
2009	941	840	478	601	825	1170	186	163	200	537	206
2010	998	899	486	688	947	1326	195	160	190	574	218
2011	933	942	549	685	1008	1249	209	147	185	560	280
2012	1067	905	595	760	1060	1258	184	163	156	528	314
2013	1036	883	576	700	1000	1278	184	179	185	495	293
2014	1247	1060	613	840	1114	1444	207	213	242	578	326
2015	1344	1125	635	1011	1071	1508	207	190	210	684	297
2016	1400	1070	628	900	829	1123	215	151	245	644	247
2017	1413	1034	606	866	822	1111	265	177	213	597	241

西元年	雲林縣	嘉義縣	屏東縣	台東縣	花蓮縣	澎湖縣	基隆市	新竹市	嘉義市	金門縣	連江縣
2008	407	379	583	130	204	47	203	130	124	17	3
2009	433	347	596	139	221	65	157	109	119	25	0
2010	446	323	627	162	203	60	174	99	99	34	1
2011	490	362	583	146	215	62	164	136	104	35	3
2012	497	380	601	150	221	61	168	123	90	30	4
2013	547	343	536	138	206	51	150	129	110	22	1
2014	555	414	612	139	223	43	171	141	127	40	4
2015	536	376	597	168	209	50	230	152	107	52	2
2016	443	350	376	122	156	54	134	134	118	32	1
2017	403	301	424	118	156	39	158	127	112	49	3

在各縣市每年工廠數方面，我們由經濟部統計處的工廠名錄中整理出 2008 年至 2017 年石油及煤製品業、基本金屬業、金屬製品業、電子零組件業，因這幾類行業工廠容易產生如文獻中所看到的致癌金屬物質或廢氣，所



## 點燃你的肺-肺癌成因之探討

以我們將以這幾類行業的各縣市工廠數做為探討的數據資料。資料如下表三所示。

表三、2008 年至 2017 年的各縣市工廠數

西元年	新北市	台北市	桃園市	台中市	台南市	高雄市	宜蘭縣	新竹縣	苗栗縣	彰化縣	南投縣
2008	167	1	113	149	81	87	9	19	15	79	4
2009	128	1	71	87	66	54	7	12	10	46	2
2010	164	4	104	139	58	81	5	24	17	77	4
2011	196	5	150	196	58	73	8	15	9	100	3
2012	269	14	148	196	73	94	13	25	14	150	6
2013	184	4	175	349	106	231	4	28	24	296	3
2014	247	1	113	340	87	147	2	25	20	155	3
2015	241	5	148	380	100	181	9	30	21	231	10
2016	251	1	150	560	123	175	12	37	39	275	5
2017	298	0	186	456	117	159	7	32	27	205	4

西元年	雲林縣	嘉義縣	屏東縣	台東縣	花蓮縣	澎湖縣	基隆市	新竹市	嘉義市	金門縣	連江縣
2008	10	5	9	1	0	0	2	21	1	0	0
2009	5	3	6	0	2	1	5	8	0	0	0
2010	11	8	5	0	0	0	5	14	1	0	0
2011	28	8	13	0	1	0	3	14	1	0	0
2012	41	8	6	0	1	0	3	19	2	0	0
2013	46	23	13	0	0	1	2	13	8	0	0
2014	20	10	17	0	0	3	3	13	1	0	0
2015	22	4	49	0	0	0	1	11	2	0	0
2016	39	20	30	2	2	0	2	13	3	0	0
2017	43	18	14	0	2	0	5	16	2	0	0

在吸菸人口比率方面，我們在政府資料開放平台上搜尋到由衛生福利部國民健康署的「國人吸菸行為調查」中 2008 年至 2017 年的 18 歲以上吸菸比率資料。資料如下表四所示。

表四、2008 年至 2017 年的 18 歲以上吸菸比率(單位：百分比)

西元年	吸菸率
2008年	21.9
2009年	20
2010年	19.8
2011年	19.1
2012年	18.7
2013年	18
2014年	16.4
2015年	17.1
2016年	15.3
2017年	14.5

我們從行政院環境保護署的環保統計查詢網的公務統計報表中查詢到「空氣品質指標分布」這份報表。在這份報表中把空氣汙染指標(PSI)大於 100 定義為空氣品質不良，因此我們的原始資料是 2008 年至 2016 年各縣市一年中空氣品質不良的比例。然後我們將所有空氣品質不良的比例取平均，並且以這個平均做為基準，把每年各縣市分成兩類，將大於平均的空氣品質不良的比例的縣市定為空氣品質較不好，然後將小於平均的空氣品質不良的比例的縣市定為空氣品質較好。

再從內政部統計處的內政部統計查詢網得到 2008 年至 2016 年的各縣市人口數，由各縣市的人口數和肺癌死亡人數計算出每年各縣市的肺癌死亡率。之後把空氣品質較不好的縣市的肺癌死亡率資料歸類到空氣品質較不好的肺癌死亡率，然後把空氣品質較好的縣市的肺癌死亡率資料歸類到空氣品質較好的肺癌死亡率。因此最後我們所呈現的資料為空氣品質較好的肺癌死亡率和空氣品質較不好的肺癌死亡率。資料如下表五所示。

表五、西元 2008 年至 2016 年的空氣品質好壞的肺癌死亡率

空氣品質較不好						
0.026	0.053	0.048	0.062	0.045	0.047	0.034
0.024	0.051	0.042	0.06	0.071	0.041	0.0082
0.023	0.044	0.043	0.056	0.072	0.038	
0.034	0.053	0.039	0.079	0.063	0.036	
0.032	0.04	0.054	0.065	0.07	0.043	
0.036	0.054	0.041	0.071	0.067	0.045	
0.026	0.052	0.039	0.067	0.072	0.039	
0.023	0.046	0.05	0.059	0.068	0.031	
0.024	0.045	0.079	0.063	0.066	0.018	
0.059	0.045	0.077	0.069	0.06	0.027	

空氣品質較好							
0.035	0.034	0.025	0.054	0.028	0.037	0.045	0.062
0.034	0.036	0.033	0.047	0.035	0.043	0.038	0.043
0.031	0.029	0.037	0.045	0.04	0.033	0.041	0.024
0.026	0.03	0.031	0.045	0.034	0.028	0.043	0.035
0.027	0.03	0.026	0.04	0.031	0.033	0.044	0.055
0.024	0.028	0.028	0.04	0.028	0.034	0.041	0.067
0.04	0.029	0.026	0.046	0.031	0.036	0.049	0.051
0.042	0.027	0.044	0.042	0.032	0.033	0.058	0.045
0.039	0.024	0.057	0.04	0.039	0.05	0.063	0.032
0.033	0.024	0.056	0.049	0.044	0.053	0.057	0.032
0.06	0.061	0.064	0.062	0.04	0.032	0.062	0.062
0.064	0.066	0.06	0.068	0.052	0.044	0.066	
0.077	0.064	0.065	0.05	0.031	0.04	0.064	
0.07	0.07	0.06	0.036	0.035	0.033	0.045	
0.069	0.056	0.052	0.062	0.033	0.024	0.026	
0.068	0.047	0.049	0.046	0.03	0.0079	0.03	

## 第二節、肺癌死亡人數與工廠數的關係

在探討肺癌死亡人數與工廠家數的關係上，由於肺癌死亡人數和工廠數皆為計量資料，因此可使用迴歸分析來了解肺癌死亡人數和工廠數的相關程度、相關方向與強度。本文以簡單線性迴歸模型  $Y = \beta_0 + \beta_1 X_1 + \varepsilon$  進行兩者關係的探討，其中應變數( $Y$ )設定為肺癌死亡人數，自變數( $X_1$ )為工廠數。

### 第三節、肺癌死亡人數與 18 歲以上吸菸率的關係

在探討肺癌死亡人數與 18 歲以上吸菸率的關係上，由於肺癌死亡人數和吸菸比率皆為計量資料，因此可使用迴歸分析來了解肺癌死亡人數和吸菸比率的相關程度、相關方向與強度。本文以簡單線性迴歸模型  $Y=\beta_0+\beta_1X_1+\varepsilon$  進行兩者關係的探討，其中應變數( $Y$ )設定為肺癌死亡人數，自變數( $X_1$ )為吸菸率。

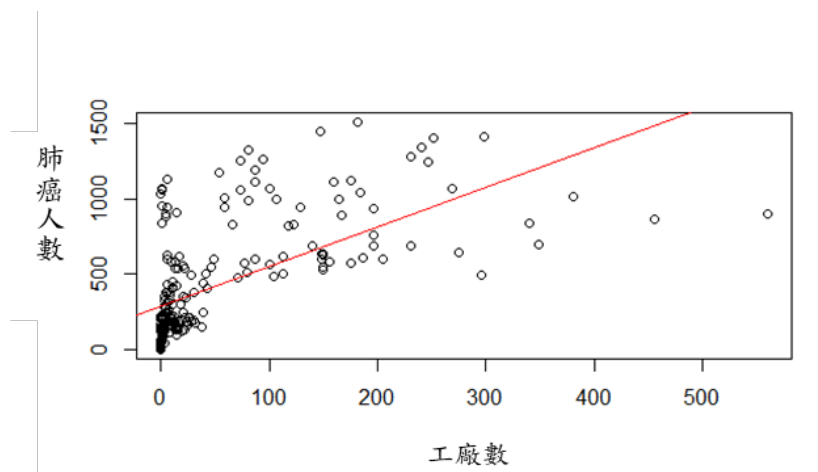
### 第四節、肺癌死亡率與空氣品質的關聯

在探討肺癌死亡率與空氣品質的關係上，我們將資料分成空氣品質較好以及空氣品質較不好兩類，設定虛無假設為空氣品質較不好的肺癌死亡率減空氣品質較好的肺癌死亡率小於等於零，對立假設為空氣較不好的肺癌死亡率減空氣較好的肺癌死亡率大於零，並利用雙獨立樣本 t 檢定做檢測，以瞭解是否空氣品質較不好時，肺癌的死亡率會較高。

## 第四章、結果分析與應用價值

### 第一節、每年肺癌死亡人數與每年工廠數的關聯(迴歸分析)

在探討每年肺癌死亡人數與工廠家數之間的關係方面，首先，以兩者的散佈圖來觀察，如圖一，可看出兩者之間呈現出正向關係的趨勢，因此，將進一步以迴歸模式來探討兩者之間的線性關係。



圖一、工廠數與肺癌死亡人數之散佈圖

本文以簡單線性迴歸模式進行探討，以工廠家數做為解釋變數，肺癌死亡人數為應變數。首先，以迴歸分析的結果來觀察工廠家數對於肺癌死亡人數的解釋能力，由 R 平方值可觀察解釋變數的解釋能力，R 平方值越接近 1 會代表解釋變數對應變數的解釋能力越好。由表六可知，以工廠數做為解釋變數來解釋肺癌死亡人數的 R 平方值有 0.387，表示有 38.7% 的解釋能力，即工廠家數能解釋肺癌死亡人數的變異之比例約為四成，顯示工廠數具有一定程度的解釋能力。

表六、工廠數對肺癌死亡人數的解釋能力

模型	R	R 平方	調整後 R 平方	估計的標準誤
1	0.622 <sup>a</sup>	0.387	0.3842	296.1

a. 解釋變數：(常數)，工廠數

在迴歸模式的係數估計方面，由表七可知，常數項( $\beta_0$ )及工廠數( $\beta_1$ )的係數估

## 點燃你的肺-肺癌成因之探討

計值皆達顯著，顯示工廠數與肺癌死亡人數具有顯著關係存在，其迴歸關係可表示為：肺癌死亡人數=286.9716+2.6259\*工廠數。因為工廠數的係數(斜率)估計值為正的，所以代表當工廠家數增加時，肺癌死亡人數也會跟著增加。

表七、工廠數對肺癌死亡人數之迴歸係數估計

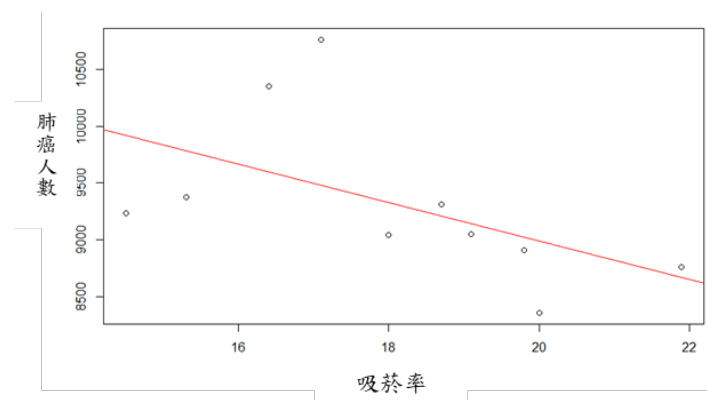
模型	非標準化係數			T	顯著性
	$\beta$	標準誤			
1	(常數)	286.9716	23.1052	12.42	<2e-16
	工廠數	2.6259	0.2238	11.73	<2e-16

此迴歸分析所使用之 R 程式碼如下：

```
setwd('C:/Users/user/Desktop')
getwd()
data<- read.csv('data.csv',header=T,sep=',')
head(data)
plot(data$factoryR,data$numberR)
abline(lm(data$numberR~data$factoryR),col="red")
s1m.model=lm(data$numberR~data$factoryR,data=data)
s1m.model
summary(s1m.model)
```

## 第二節、全國每年肺癌人數與全國 18 歲以上吸菸比率的關聯(迴歸分析)

在探討每年肺癌死亡人數與吸菸比率之間的關係方面，首先，以兩者的散佈圖來觀察，如圖二，可看出兩者之間呈現出負向關係的趨勢，因此，將進一步以迴歸模式來探討兩者之間的線性關係。



圖二、吸菸率與肺癌死亡人數的關係之散佈圖

本文以簡單線性迴歸模式進行探討，以吸菸比率做為解釋變數，肺癌死亡人數為應變數。首先，以迴歸分析的結果來觀察吸菸比率對於肺癌死亡人數的解釋能力，由 R 平方值可觀察解釋變數的解釋能力，R 平方值越接近 1 會代表解釋變數對應變數的解釋能力越好。由表八可知，以吸菸比率做為解釋變數來解釋肺癌死亡人數的 R 平方值有 0.284，表示有 28.4% 的解釋能力，即吸菸比率能解釋肺癌死亡人數的變異之比例約為三成，顯示吸菸比率對於肺癌死亡人數變化的解釋能力仍有待提升。

表八、吸菸率對肺癌死亡人數的解釋能力

模型	R	R 平方	調整後 R 平方	估計的標準誤
2	0.533 <sup>a</sup>	0.284	0.195	94.7

a. 解釋變數：（常數），吸菸率

在迴歸模式的係數估計方面，由表九可知，常數項( $\beta_0$ )的估計值達顯著而吸菸比率( $\beta_1$ )的係數估計值未達顯著，顯示吸菸比率與肺癌死亡人數不具有線性關係，表示吸菸比率對於肺癌死亡人數的變化不具有顯著的解釋能力。

表九、吸菸比率對肺癌死亡人數之迴歸係數估計

模型	非標準化係數				
	$\beta$	標準誤	T	顯著性	
2	(常數)	12368.6	1724.4	7.173	9.5e-05
	吸菸比率	-168.9	94.7	-1.783	0.112

此迴歸分析所使用之 R 程式碼如下：

```
setwd('C:/Users/user/Desktop')
getwd()
work<- read.csv('work.csv',header=T,sep=',')
head(work)
plot(work$rateR,work$numberR)
abline(lm(work$numberR~work$rateR),col="red")
slm.model=lm(work$numberR~work$rateR,data=work)
slm.model
summary(slm.model)
```

### 第三節、肺癌死亡率與空氣品質的關聯(雙樣本平均數差異 t 檢定)

本文為了進一步了解空氣品質對於肺癌死亡人數的影響，利用空氣汙染指標區分空氣品質的好壞，並透過假設檢定方式來檢測在空氣品質好與不好的情況下，肺癌死亡率之差異。本研究以雙獨立樣本 t 檢定來進行兩種空氣品質情況下的肺癌死亡率差異分析，此假設檢定之虛無假設與對立假設如下：

$H_0$ : 空氣較不好的肺癌死亡率  $\leq$  空氣較好的肺癌死亡率

$H_a$ : 空氣較不好的肺癌死亡率  $>$  空氣較好的肺癌死亡率

由表十的分析結果可知，t 檢定的結果顯示  $p\text{-value}=0.01605 < \alpha = 0.05$ ，即表示有足夠證據可拒絕虛無假設  $H_0$ ，可知在不同空氣品質之下，肺癌死亡率具有顯著差異，空氣品質較不好的情況下。肺癌死亡率大於空氣品質較好的情況，所以空氣品質好壞會影響肺癌死亡率。



表十、不同空氣品質下的肺癌死亡人數平均數差異獨立樣本 t 檢定

獨立樣本 t 檢定			
肺癌人數	t	自由度	顯著性(單尾)
採用相等變異數	2.1718	105.68	0.01605

此獨立樣本 t 檢定所使用之 R 程式碼如下：

```
setwd('C:/Users/user/Desktop')
getwd()
data2<- read.csv('data2.csv',header=T,sep=',')
head(data2)

attach(data2)
t.test(x=Badday, y=Notbadday,alternative='greater', conf.level=0.95)
detach(data2)
```

本研究的資料整理與統計分析，皆以 Excel 與 Rstudio 統計軟體進行操作。

#### 第四節、應用價值

我們可以將這次的研究結果，提供給一些醫療機構，讓他們能夠朝著這些與肺癌相關的原因來加以防範及進行宣導，降低未來肺癌的發生率。更甚提供給管理工業相關的機構，希望藉以管控工廠數或是加強督促工廠的部分。期許讓更多單位合作一起降低肺癌罹癌率。即使我們做出來的結果顯示，吸菸與肺癌死亡人數可能無關，但不代表對健康沒有危害，根據高雄縣政府衛生局表示，例如：呼吸道的免疫力下降、胃潰瘍、骨質疏鬆、女性不易受孕以及影響女性荷爾蒙使得更年期提早到來等問題，所以呼籲大眾盡量不要吸菸。

## 第五章、結論與研究限制

### 第一節、結論

本文利用迴歸模式分析每年肺癌人數與每年工廠數之間的關係，所得的結果從散佈圖得知，工廠數的多寡會影響罹患肺癌的人數，所以表示若某地區的工廠數增加，那麼該地區的肺癌死亡人數可能也會因此上升，與我們以往的認知是相同的。另外，我們也使用迴歸模式分析全國每年肺癌死亡人數與全國18歲以上吸菸比率之間的關聯，由分析結果得知，肺癌死亡人數與吸菸比率並無線性關係，所以吸菸比率增加或減少可能不會影響肺癌死亡人數的增減，因此我們得到的結果顯示吸菸比率與罹患肺癌無顯著的線性關係。最後，我們使用獨立樣本t檢定分析空氣品質好與不好是否會造成肺癌死亡率具有差異，從結果得出，空氣品質的好、壞的確會造成肺癌死亡率具有顯著差異。

### 第二節、研究限制

首先，在這次的報告中，我們在蒐集資料時因為資料很零散所以花了很多時間才找到完整的資料，或著是因為資料不是我們所要的，所以只能將找到的資料重新整理成我們要的。其次我們找的資料中礙於縣市的合併，我們必須將縣市合併前後的資料統整，然後我們在搜尋肺癌死亡人數時也發現，2007年之前只有全部癌症的死亡人數，而在2007年之後才有各種癌症的死亡人數，所以肺癌死亡人數只有2008年至2017年的資料。最後，在空氣品質資料方面，由於空氣污染的標準在2016/12/31之前是以PSI來作為指標，而在2017/01/01之後改成以AQI作為指標，為了資料的一致性，我們在探討肺癌死亡率和空氣品質的關聯時只討論2008年至2016年期間的資料。由於以上的種種限制，我們能夠取得的資料可能不夠多也不盡完整，因此也造成我們分析的結果有再精進的空間。

## 第六章、團隊合作之學習與成長

在這次課程的期末報告中，我們更加熟悉及熟練R語言的操作，而且也能將大一所學的基礎統計知識應用於此報告上。透過這個報告除了能學習到專業知識，更可以幫助我們釐清我們想解決的問題。我們也藉由這個機會去印證，新聞上的資訊是否屬實，畢竟數據也是會隨著時間而有所變動。雖然現在網路資訊相當發達，我們能很容易搜尋到資料，但是要找到符合我們研究所需的資料是一件困難的事，訓練我們搜尋資料的能力，讓我們在未來研究中更加上手。



## 第七章、參考文獻與資料來源

1. 衛生福利部統計處。取自 <https://dep.mohw.gov.tw/DOS/lp-1861-113.html>
2. 經濟部統計處。取自 <https://dmz26.moea.gov.tw/GMWeb/investigate/InvestigateFactory.aspx>
3. 行政院環境保護署。取自 <https://stat.epa.gov.tw/>
4. 臺北市立聯合醫院癌症防治研究發展中心癌症資訊網。取自 <http://website.tp.ech.gov.tw/cancer/%E7%99%8C%E7%97%87%E9%A0%90%E9%98%B2%E8%88%87%E7%AF%A9%E6%AA%A2/%E8%82%BA%E7%99%8C%E7%B0%A1%E4%BB%8B.html>
5. 新光醫療財團法人-新光吳火獅紀念醫院胸腔外科。取自 <http://www.skh.org.tw/ts/service08.asp>
6. 內政部統計處。取自 <http://statis.moj.gov.tw/micst/stmain.jsp?sys=100>
7. 行政院性別平等會。取自 [https://www.gender ey.gov.tw/gecdb/Stat\\_Statistics\\_DetailData.aspx?sn=ORVNSDIGVFTYgkN%2B7bdIyA%3D%3D](https://www.gender ey.gov.tw/gecdb/Stat_Statistics_DetailData.aspx?sn=ORVNSDIGVFTYgkN%2B7bdIyA%3D%3D)
8. 楊智仁 (2015)。「罹患肺癌的可能危險因子」。高醫醫訊。取自 <http://www.kmu.edu.tw/www/kmcj/data/10407/4.html>
9. Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018) Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries, *CA CANCER J CLIN*, **68**, 394–424.