

中級マイクロデータサイエンス Problem set2

学籍番号：2125095

氏名：佐々木 棕

コードのURL

<https://github.com/Ryo-Sasaki-xxx/problemset>にプッシュしてあります。

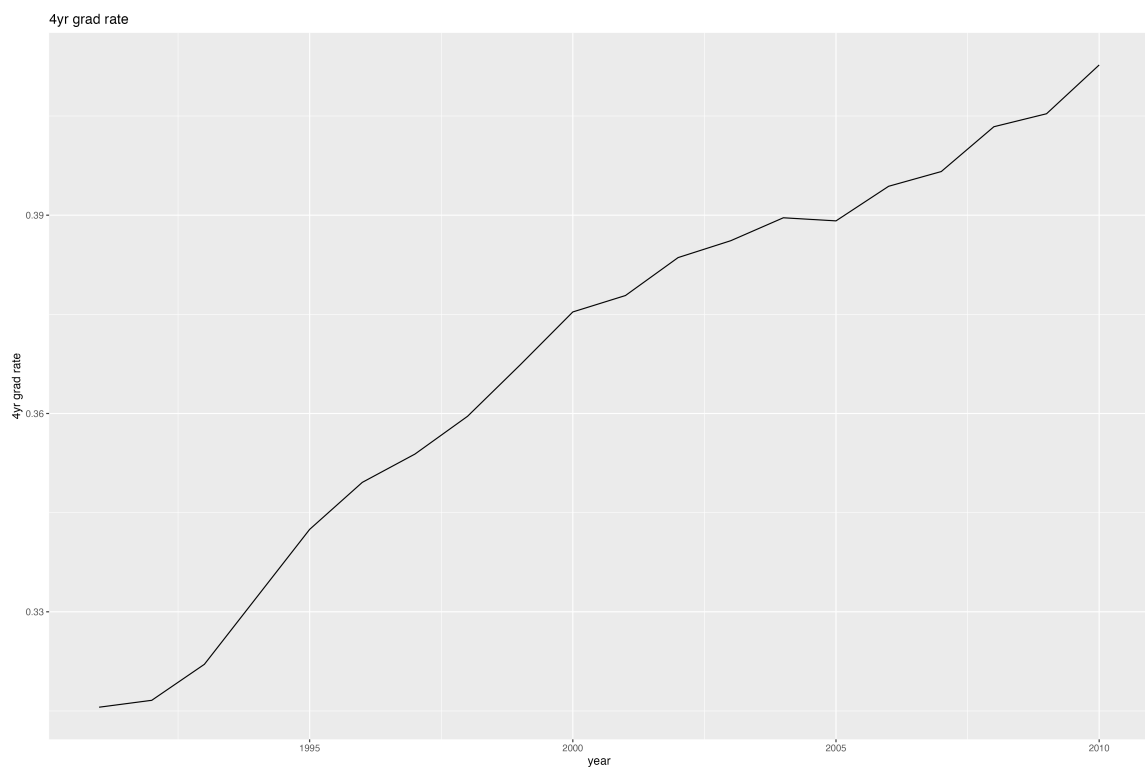
(a)記述統計

1. 問題背景などを知る上で役に立つ記述統計を作成し、内容について議論しなさい

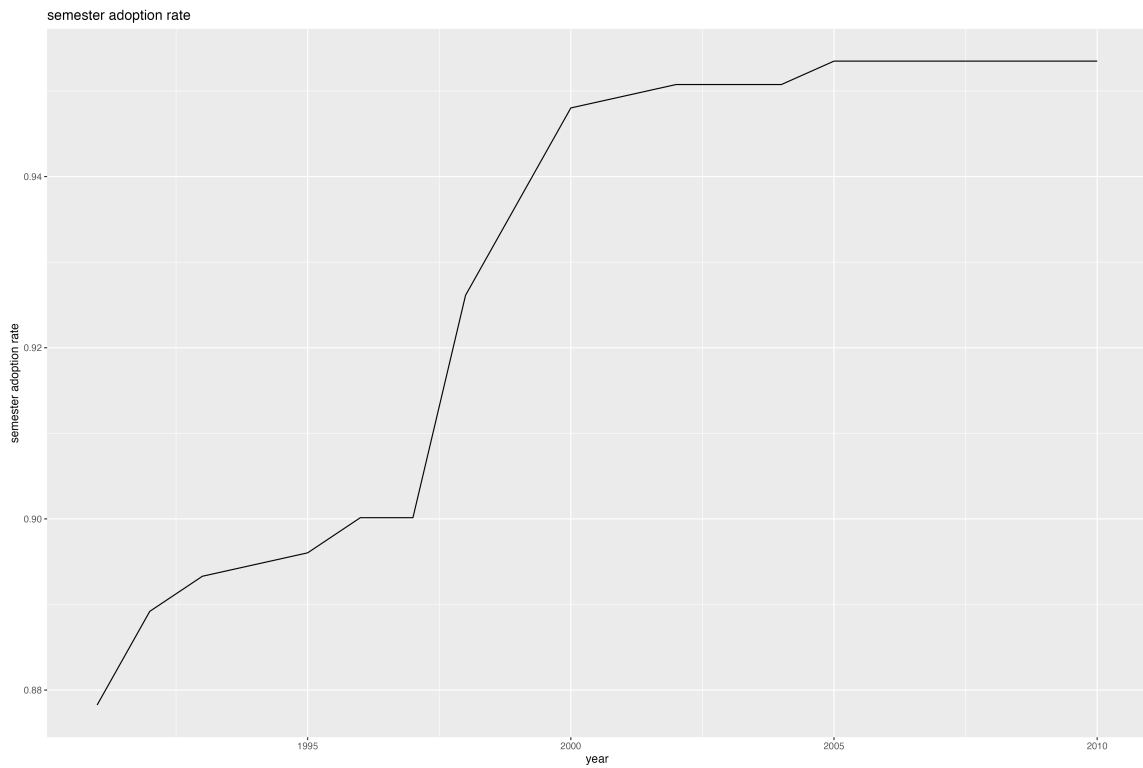
Table			
Characteristic	Overall, N = 13,889 [†]	changed to semester, N = 1,045 [†]	still quarter, N = 12,844 [†]
cohort size	1,099 (1,183)	1,279 (1,320)	1,085 (1,170)
women's cohort size	599 (629)	686 (724)	592 (620)
men's cohort size	500 (571)	593 (620)	492 (566)
women's 4yr grad rate	0.41 (0.23)	0.32 (0.20)	0.42 (0.23)
Unknown	24	0	24
men's 4yr grad rate	0.32 (0.23)	0.22 (0.18)	0.33 (0.23)
Unknown	65	0	65
women's cohort size(%)	0.56 (0.10)	0.54 (0.10)	0.56 (0.10)
white cohort size(%)	0.71 (0.24)	0.72 (0.24)	0.71 (0.24)
[†] Mean (SD)			

セメスター制に移行した大学と移行していない大学とを比較すると、男女ともに移行していない大学の方が4年卒業率が高いことがわかる。さらに4年卒業率を男性と女性で比較すると、女性の方が高いことがわかる。また話題から逸れるかもしれないが、アメリカの大学では女性の方が、男性よりも多いことに少し驚いた。

2. 4年卒業率の平均推移を計算し、図で示しなさい

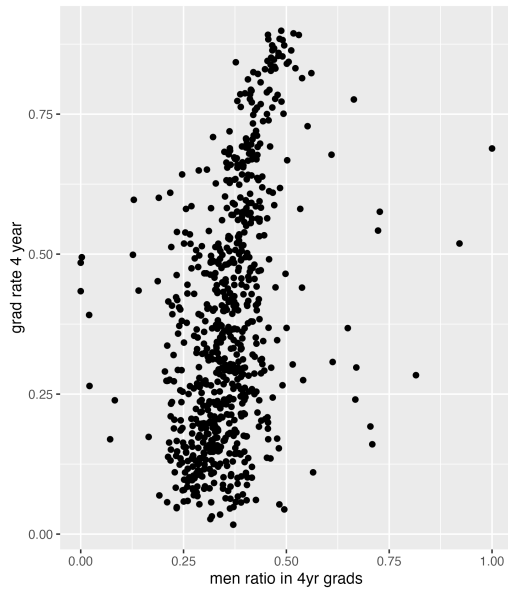


3. semester導入率を計算し、図で示しなさい

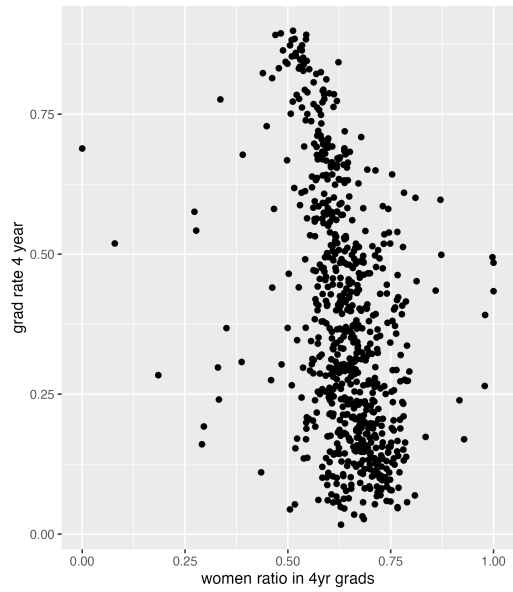


4. 変数に処理を加えた上で、以下の散布図を作成しなさい。また、重要だと考える結果について議論しなさい

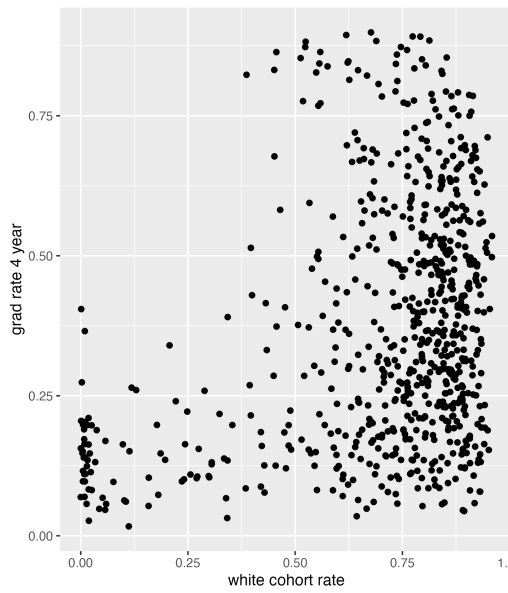
4年卒業率と男子学生比率



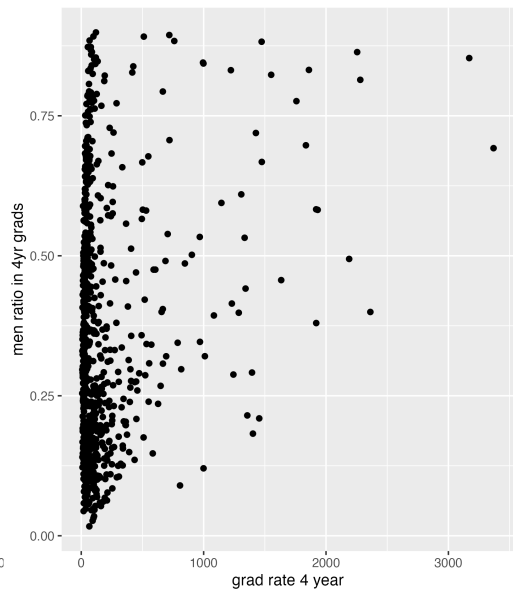
4年卒業率と女子学生比率



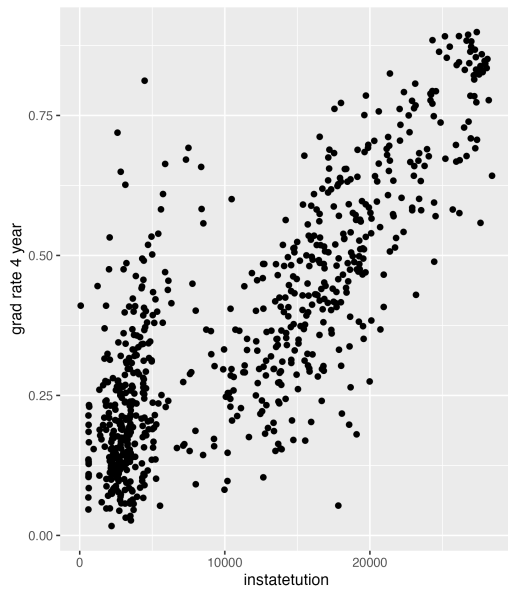
4年卒業率と白人学生割合



4年卒業率と年間運営コスト



4年卒業率と学費



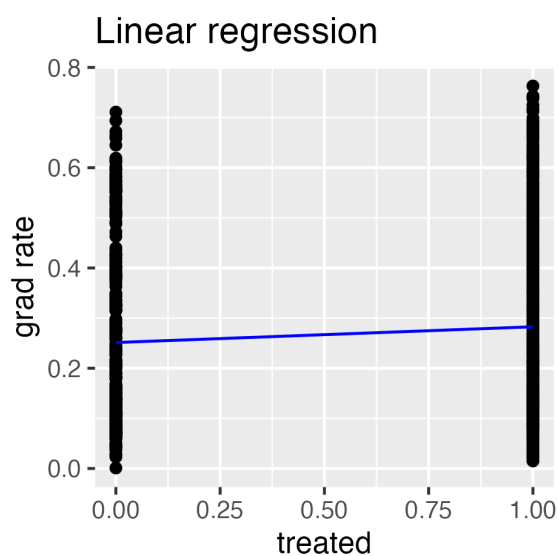
4年卒業率と学費の散布図は大変特徴的である。この図から学費と4年卒業率には強い相関関係があることがわかる。私の個人的な考察であるが、高い学費を払える家庭の子供はアルバイトに割く時間を短くできることや、自分で家事をしなくてもよいことなど、学業に集中できる環境下であるからではないだろうか。

(b)回帰分析

1.3つの変数を作成しなさい

problemset2_based/02_analysis/code/regression.Rが該当するスクリプトである。

2.式を推定し、結果について議論しなさい



Characteristic	N = 2 ¹
term	
(Intercept)	1 (50%)
treated	1 (50%)
estimate	
0.0314272891349865	1 (50%)
0.251244372990354	1 (50%)
std.error	
0.0103895832203053	1 (50%)
0.0123967613699816	1 (50%)
statistic	
2.5351209236863	1 (50%)
24.1823341382283	1 (50%)
p.value	
6.33323848094196e-103	1 (50%)
0.0113860385037467	1 (50%)
¹ n (%)	

結果的には以下のような式と推定できた。

$$Y_{sk} = 0.251 + 0.031treated_{sk} + \varepsilon_{sk}$$

よってセメスター制に移行すると3.1%ほど4年卒業率が上がることが予想される。

3.数式の問題点を指摘しなさい

説明変数が一つだけなので、セメスター制に移行したという要因だけでは必ずしも3.1%ほど4年卒業率が上がると説明できないことは問題だと考える。他の要因によってこのような結論になっただけなのかもしれない。