初歩的な統計講座（午後・演習編）

山本義郎（東海大学）
藤野友和（福岡女子大学）

データについて

- ExcelIR演習_PTest.xlsx(入学時のデータ)
  - プレースメントテスト英語(100点満点)
  - プレースメントテスト数学(100点満点)
  - 入学種別（AO入試、推薦入試、一般入試、センター入試）
- 成績.xlsx(1年次の成績データ)
  - 1セメスター
    - GPA（S:4点, A:3点, B:2点, C:1点, 不合格:0点）
    - 修得単位数（上限 24単位/1セメスター）
  - 2セメスター
    - 学期GPA、通算GPA
    - 学期単位数、通算単位数

Vlookupを用いたデータの参照

- １年次成績.xlsx　に「新しいシート」を追加し、「全体」シートの内容をコピー＆貼り付け
- 「1セメGPA」と変数名を入力し F2 セルに式 =VLOOKUP(A2,'[ExcelIR演習_成績.xlsx]1セメスター'!$A$2:$C$248,2,FALSE)

同様に、1セメ単位数、2セメ学期GPA、2セメ学期単位数、学科名を作成する

配布資料に
「 ExcelIR演習_１年次成績.xlsx 」としてこの結果があります。
うまくできなかった方は、このファイルを開いて、次の演習に取り組んでください。

【演習１】分析データの準備

２つのファイルから、各学生の入学時と1年次の成績を結合して、成績について分析できる準備をする。

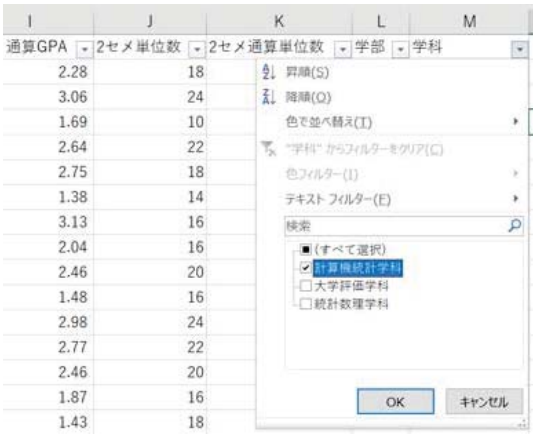
（留意点）

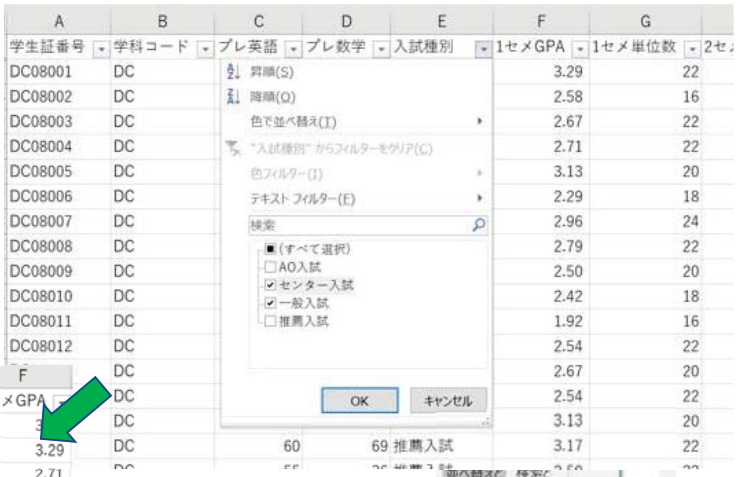
- 入学時データ()をコピーし「１年次成績.xlsx」と名前を変更
- 値の参照表示（vlookup）
- 結合時の留意点（変数名：各セメスターのGPA、単位数）

【演習２】フィルターとソート

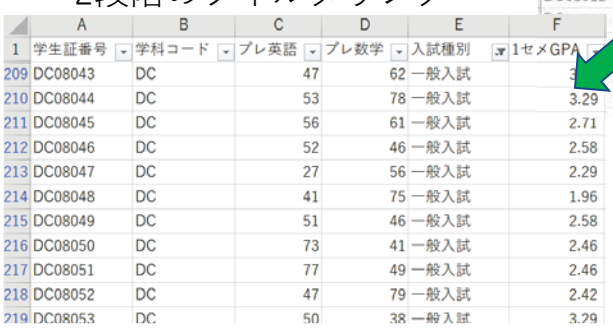
1. フィルター
   - 「計算機統計学科」の「一般入試」と「センター入試」の学生だけをファルターで抽出
2. ソート
   1. プレースメントテストの数学の得点の高い順にソートする
   2. 入試種別ごとに、プレースメントテストの数学の得点の高い順にソートする

フィルター



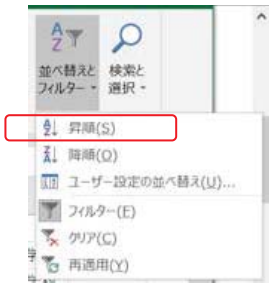


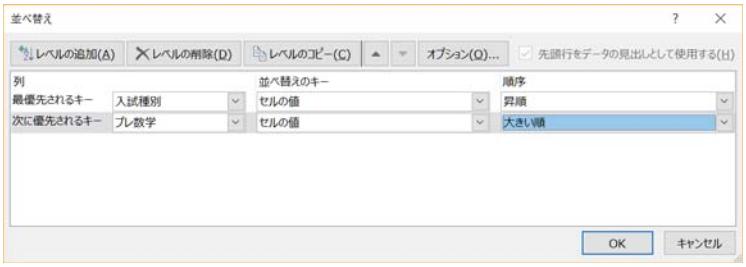
2段階のフィルタリング



フィルタ解除

ソート





昇順：値の小さい順
降順：値の大きい順

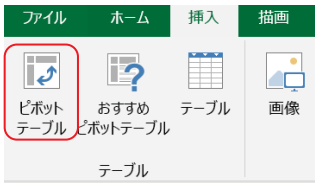
2つ以上の変数に関してソート（DCのみ）

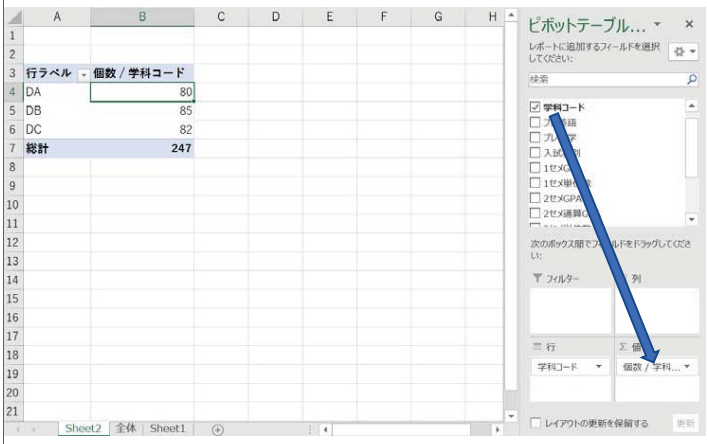
	A	B	C	D	E	F	G
1	学生証番号	学科コード	プレ英語	プレ数学	入試種別	1セメGPA	1セメ単位数
167	DC08007	DC	67	64	AO入試	2.96	24
168	DC08003	DC	49	51	AO入試	2.67	22
169	DC08001	DC	46	49	AO入試	3.29	22
170	DC08004	DC	38	43	AO入試	2.71	22
171	DC08008	DC	39	41	AO入試	2.79	22
172	DC08011	DC	64	40	AO入試	1.92	16
173	DC08010	DC	29	38	AO入試	2.42	18
174	DC08012	DC	31	38	AO入試	2.54	22
175	DC08002	DC	56	35	AO入試	2.58	16
176	DC08005	DC	51	33	AO入試	3.13	20
177	DC08006	DC	43	27	AO入試	2.29	18
178	DC08009	DC	51	23	AO入試	2.50	20
179	DC08072	DC	47	67	センター入試	2.29	22
180	DC08077	DC	54	66	センター入試	2.92	24
181	DC08076	DC	56	64	センター入試	2.67	16
182	DC08074	DC	69	62	センター入試	2.46	20
183	DC08079	DC	56	56	センター入試	2.29	20

【演習３】集計（ピボットテーブル）

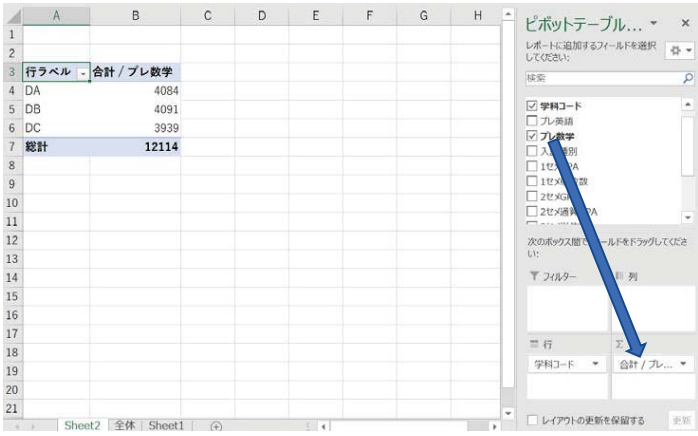
1. 質的変数の集計
   1. 学科ごとの学生数を集計する（単純集計）
   2. 学科ごとの入試種別ごとの人数を集計する（クロス集計）
2. 量的変数の要約
   1. 学科ごとにプレースメントテストの数学の点数の平均値、標準偏差を求める

ピボットテーブル

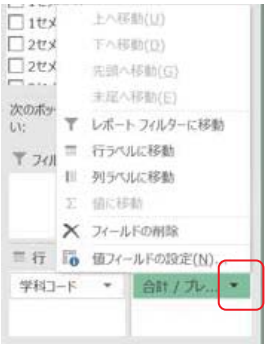




質的変数のΣ値は個数



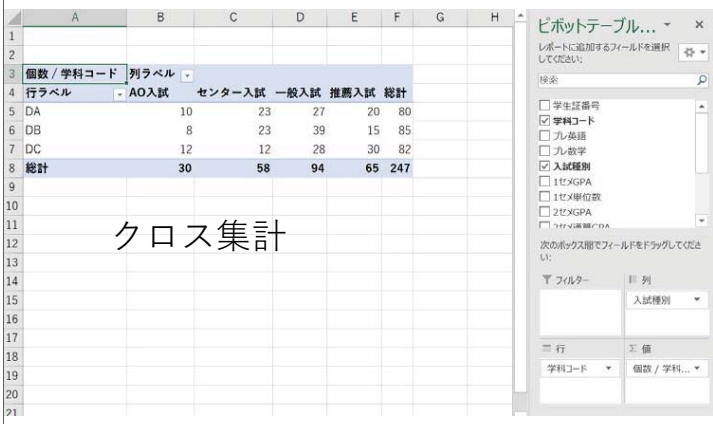
量的変数のΣ値は合計



▼クリックし
「値フィールドの設定」
で、平均などに変更可能

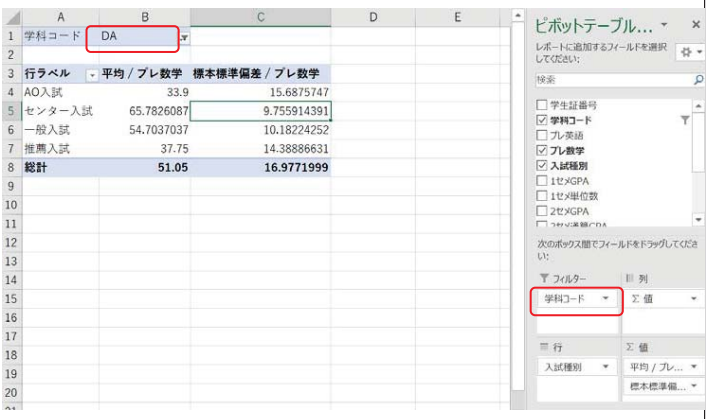
ピボットテーブル(2)

質的変数のΣ値（個数）



クロス集計

量的変数のΣ値（平均）



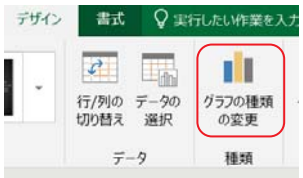
フィルターに質的変数を入れることで
フィルタリングも可能
量的変数の層別の要約は複数も可能

【演習４】グラフによる可視化

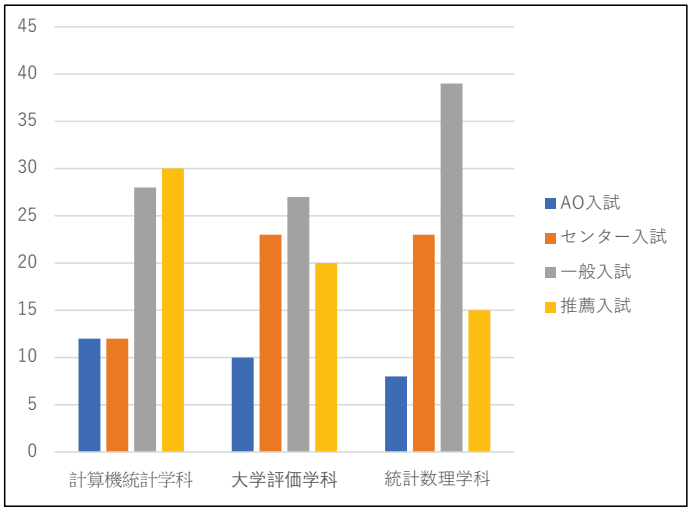
1. 質的変数の集計を棒グラフ
   1. 学科ごとの学生数の縦棒グラフ
   2. 学科ごとの入試種別ごとの人数の複合棒グラフ、積み上げ棒グラフ、帯グラフ（１００％積み上げ横棒グラフ）

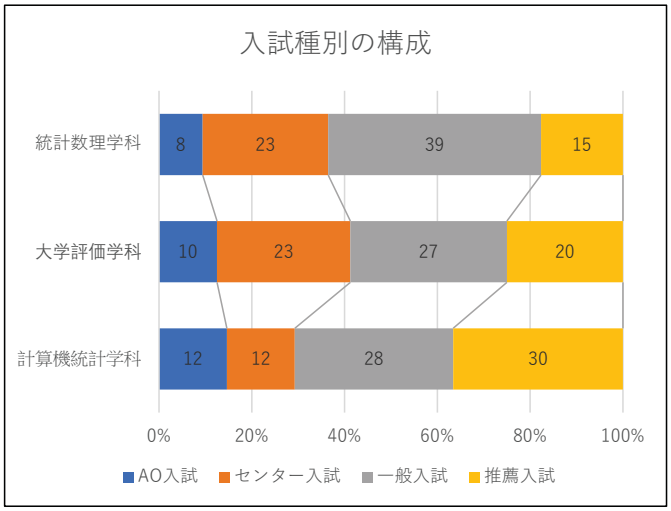
2. 量的変数の要約
   1. １セメスターのGPAのヒストグラムと、学科ごとの１セメスターのGPAのヒストグラム
   2. 学科ごとの１セメスターのGPAの箱ひげ図
   3. 学科ごとのプレースメントテストの数学の点数と１セメスターのGPAの散布図（入試種別を異なるシンボルで）

グラフ（棒グラフ、帯グラフ）

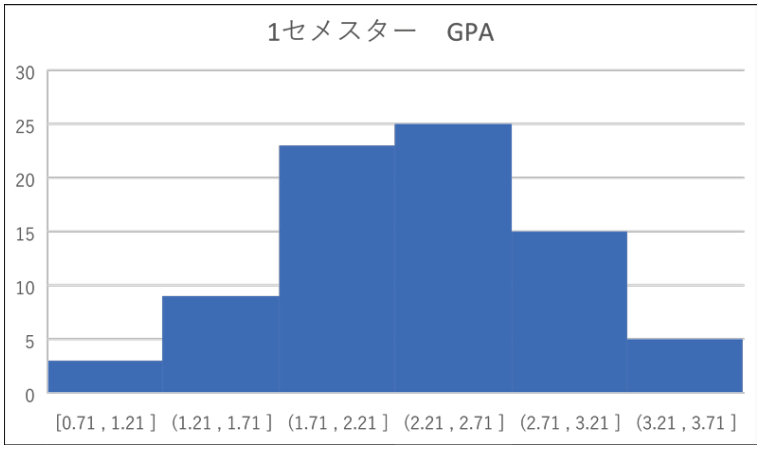


ピボットテーブルのクロス集計から作成

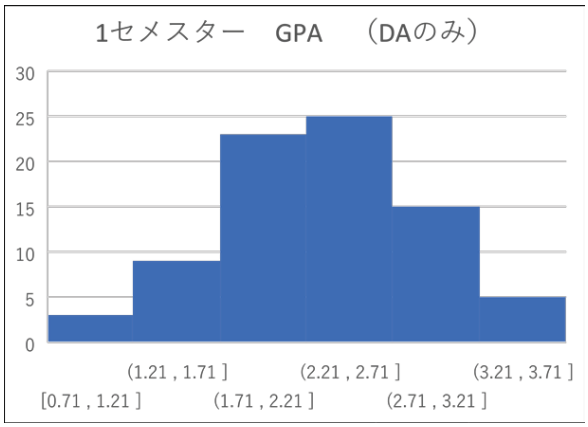




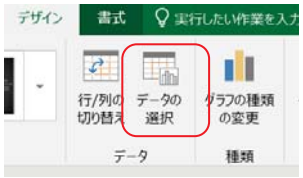
ヒストグラム

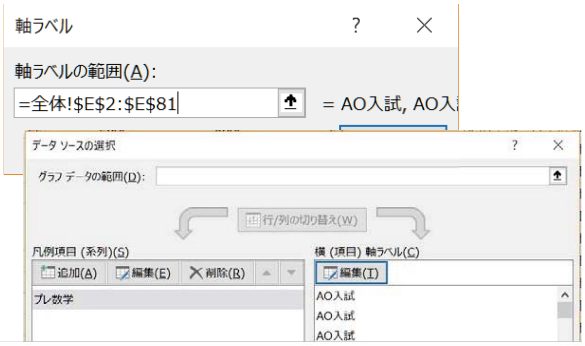


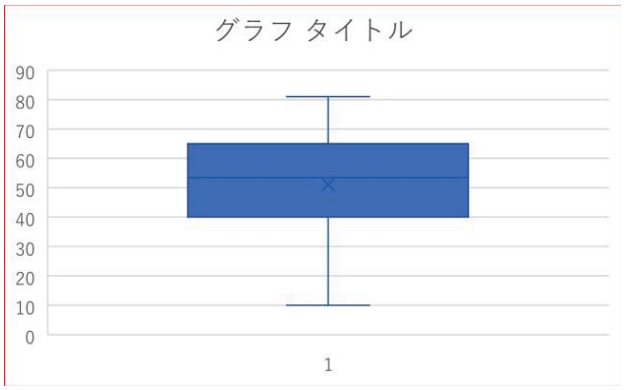
フィルターを使うと
フィルタリングされたグラフに

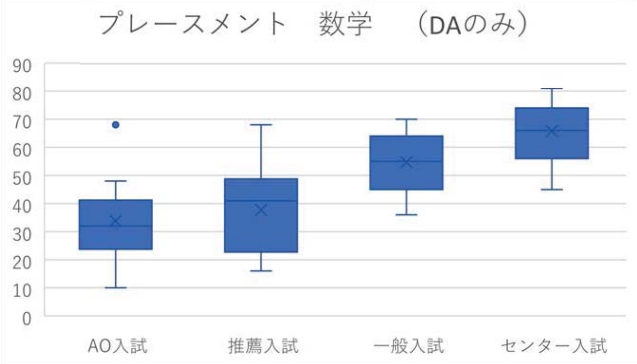


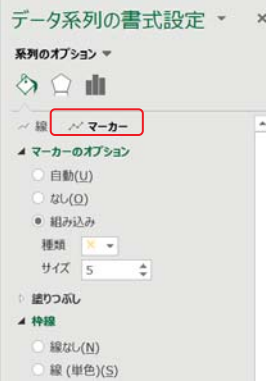
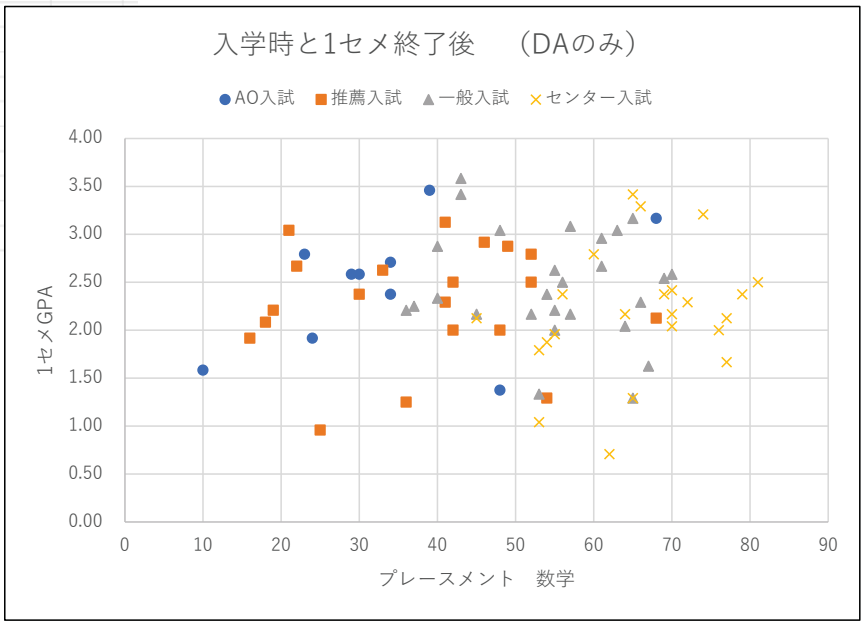
ボックスプロット（箱ひげ図）









散布図（グループ別）

	A	B	C	D	E	F	G	H
1	学生証番号	学科コード	プレ数学	入試種別	AO入試	推薦入試		
2	DA07001	DA	29	AO入試	2.58			
3	DA07002	DA	68	AO入試	3.17			
4	DA07003	DA	24	AO入試	1.92			
5	DA07004	DA	34	AO入試	2.71			
6	DA07005	DA	23	AO入試	2.79			
7	DA07006	DA	48	AO入試	1.38			
8	DA07007	DA	39	AO入試	3.46			
9	DA07008	DA	34	AO入試	2.38			
10	DA07009	DA	30	AO入試	2.58			
11	DA07010	DA	10	AO入試	1.58			
12	DA07011	DA	46	推薦入試		2.92		
13	DA07012	DA	21	推薦入試		3.04		
14	DA07013	DA	42	推薦入試		2.50		
15	DA07014	DA	16	推薦入試		1.92		
16	DA07015	DA	36	推薦入試		1.25		
17	DA07016	DA	18	推薦入試		2.08		
18	DA07017	DA	22	推薦入試		2.67