

Desaio - 07

Iuri Santos Oliveira

2025-09-23

Atividade

1. Baixe o arquivo `uwmadison.sqlite3`. Conecte-se a ele usando o pacote `RSQLite`, armazenando a conexão em uma variável `conn`.

```
library(RSQLite)
conn <- dbConnect(SQLite(), dbname = 'database.sqlite3')
dbListTables(conn)
```

```
## [1] "course_offerings"      "courses"               "grade_distributions"
## [4] "instructors"           "rooms"                  "schedules"
## [7] "sections"              "subject_memberships"    "subjects"
## [10] "teachings"
```

2. Quem são e quantos são os professores que lecionaram disciplinas cujo tópico era estatística (`subjects.abbreviation='STAT'`);

```
sql <- "
SELECT DISTINCT
    i.name as professor_nome,
    COUNT(DISTINCT s.uuid) as total_disciplinas_stat
FROM instructors i
JOIN teachings t ON i.id = t.instructor_id
JOIN sections s ON t.section_uuid = s.uuid
JOIN course_offerings co ON s.course_offering_uuid = co.uuid
JOIN subject_memberships sm ON co.uuid = sm.course_offering_uuid
JOIN subjects sub ON sm.subject_code = sub.code
WHERE sub.abbreviation = 'STAT'
GROUP BY i.name
ORDER BY total_disciplinas_stat DESC;"

# Executar a consulta
professores <- dbGetQuery(conn, sql)

professores
```

##	professor_nome	total_disciplinas_stat
## 1	JOHN GILLET	68
## 2	ERIK NORDHEIM	65
## 3	ZHENGJUN ZHANG	62
## 4	WEI-YIN LOH	56
## 5	QIAN ZHIGUANG	49
## 6	ZHANG CHUNMING	48
## 7	JUN SHAO	46
## 8	KAM-WAH TSUI	43

## 9	BRET LARGET	43
## 10	RICHARD J. CHAPPELL	42
## 11	ISMOR FISCHER	41
## 12	ROBERT WARDROP	40
## 13	KARL ROHE	39
## 14	DONALD PORTER	39
## 15	MICHAEL NEWTON	38
## 16	CECILE ANE	38
## 17	SUNDUZ KELES	37
## 18	BRIAN YANDELL	37
## 19	YAZHEN WANG	36
## 20	SHENG WANG	35
## 21	JUN ZHU	33
## 22	RUOSI GUO	32
## 23	GRACE WAHBA	31
## 24	BRET HANLON	31
## 25	CHRISTINA M. KENDZIORSKI	30
## 26	SIJIAN WANG	29
## 27	XINYU SONG	28
## 28	TONG LI	28
## 29	SHENG ZHANG	27
## 30	JIE ZHANG	26
## 31	MURRAY CLAYTON	25
## 32	JU HEE CHO	25
## 33	ZUOFENG SHANG	24
## 34	NICHOLAS STEPHEN KEULER	24
## 35	ZHUANG WU	23
## 36	YI CHAI	23
## 37	XIN ZHANG	23
## 38	QUOC TRAN	23
## 39	GARVESH RASKUTTI	22
## 40	DOUGLAS M. BATES	22
## 41	XIAO GUO	21
## 42	SHENGJI JIA	21
## 43	KYUNGMAAN KIM	21
## 44	JINGLAN LI	21
## 45	HYEBIN SONG	21
## 46	DUZHE WANG	21
## 47	XIAOMAO LI	20
## 48	WENZHI CAO	20
## 49	TIMO SEPPALAINEN	20
## 50	JINGCI MENG	20
## 51	DONGGYU KIM	20
## 52	YOUNGDEOK HWANG	19
## 53	YAN CHEN	19
## 54	SONG WANG	19
## 55	RYAN ZEA	19
## 56	RONALD GANGNON	19
## 57	QUEFENG LI	19
## 58	NING FAN	19
## 59	MOO K CHUNG	19
## 60	MINJING TAO	19
## 61	LONG PHAN	19
## 62	LILUN DU	19

## 63	JINGJIANG PENG	19
## 64	JIAJIE CHEN	19
## 65	FAN YANG	19
## 66	YI LI	18
## 67	TIMOTHY IDOWU	18
## 68	QI TANG	18
## 69	JOHN DAVIS	18
## 70	CHANHAN HSU	18
## 71	YUQING XU	17
## 72	YAOYAO XU	17
## 73	SANGBUM CHOI	17
## 74	PERLA REYES	17
## 75	HAO TENG	17
## 76	YAOGUO XIE	16
## 77	WENWEN ZHANG	16
## 78	KARL BROMAN	16
## 79	JILI WANG	16
## 80	HAOYANG FAN	16
## 81	HAN CHEN	16
## 82	GUN WOONG PARK	16
## 83	GUILHERME VIEIRA NUNES LUDWIG	16
## 84	BIN ZHU	16
## 85	BENEDEK VALKO	16
## 86	AKICHIKA OZEKI	16
## 87	ZIFENG ZHAO	15
## 88	YANG ZHAO	15
## 89	XIPEI YANG	15
## 90	KJELL DOKSUM	15
## 91	JIANCHANG HU	15
## 92	CRYSTAL CHEN	15
## 93	CHEN CHENG	15
## 94	XINXIN YU	14
## 95	XIN LI	14
## 96	WESLEY CHANG	14
## 97	STEPHEN WRIGHT	14
## 98	SHIXUE LIU	14
## 99	LIE XIONG	14
## 100	HAO ZHENG	14
## 101	HAO CHEN	14
## 102	GONZALO CONTADOR	14
## 103	DEYUAN JIANG	14
## 104	DAVID GRIFFEATH	14
## 105	BINGYING XIE	14
## 106	YUANZHI LI	13
## 107	YU MENGANG	13
## 108	SEULKEE YUN	13
## 109	SEBASTIEN ROCH	13
## 110	QING LI	13
## 111	PAUL RATHOUZ	13
## 112	MIN NIU	13
## 113	LIAM JOHNSTON	13
## 114	KEVIN PACKARD	13
## 115	JIWEI ZHAO	13
## 116	GINA BENNINGER	13

## 117	CUICUI QI	13
## 118	BO YANG	13
## 119	YUAN WANG	12
## 120	YOURAN QI	12
## 121	YILIN ZHANG	12
## 122	TRAM TA	12
## 123	SIJING LI	12
## 124	PEIGEN ZHOU	12
## 125	MING XIE	12
## 126	LU YANG	12
## 127	KUNLING HUANG	12
## 128	JUN LI	12
## 129	JIE SONG	12
## 130	JENNIFER BIRSTLER	12
## 131	HAODA FU	12
## 132	HAO ZHOU	12
## 133	FAN GAO	12
## 134	CHENXI LI	12
## 135	ANQI SHI	12
## 136	YUAN LI	11
## 137	YOUNG LEE	11
## 138	YONGFENG WU	11
## 139	THOMAS G. KURTZ	11
## 140	TAO YU	11
## 141	SEUNGBONG HAN	11
## 142	SEHO PARK	11
## 143	QIONG ZHANG	11
## 144	MUHONG GAO	11
## 145	JIALE XU	11
## 146	DAVID ANDERSON	11
## 147	CUIZE HAN	11
## 148	CHEN JING	11
## 149	XU XU	10
## 150	VICTOR LUO	10
## 151	SOHEIL SADEGHI	10
## 152	NORBERT BINKIEWICZ	10
## 153	MICHAEL FERRIS	10
## 154	JUN ZHANG	10
## 155	JARED BROWN	10
## 156	ERICA LEE DEADMAN	10
## 157	DEREK BEAN	10
## 158	ZHONGJIE YU	9
## 159	YONGSU LEE	9
## 160	YI LIU	9
## 161	XIAO NIE	9
## 162	SHIZHEN WANG	9
## 163	RUI CHEN	9
## 164	REBECCA KOSCIK	9
## 165	QIURONG CUI	9
## 166	MENG SONG	9
## 167	JUN YIN	9
## 168	JOSEPH DEUTSCH	9
## 169	FREDERICK BOEHM	9
## 170	CHENLIANG XU	9

## 171	BIN DAI	9
## 172	YUAN JIANG	8
## 173	YOUNG MIN PARK	8
## 174	YING ZHANG	8
## 175	XIWEN MA	8
## 176	XIUYU MA	8
## 177	THOMAS COOK	8
## 178	SOKOL VAKO	8
## 179	SHANG WU	8
## 180	RUI TANG	8
## 181	RICHARD A. JOHNSON	8
## 182	MING YUAN	8
## 183	LUXI CAO	8
## 184	LUWAN ZHANG	8
## 185	HUIKUN ZHANG	8
## 186	HEATHER MARIE BRAZEAU	8
## 187	GREGORY SHINAULT	8
## 188	YUCHEN ZHOU	7
## 189	XIRAN WANG	7
## 190	TZU HSIANG HUNG	7
## 191	TIEN VO	7
## 192	THEVAASIINEN CHANDERENG	7
## 193	PAUL M TERWILLIGER	7
## 194	MARY LINDSTROM	7
## 195	LILI LAN	7
## 196	DUY NGUYEN	7
## 197	DONG XIA	7
## 198	CLAUDIA SOLIS LEMUS	7
## 199	CHAOYANG YU	7
## 200	BROOK LUERS	7
## 201	BOWEN HU	7
## 202	ADIN-CRISTIAN ANDREI	7
## 203	ABIGAIL BENZINE	7
## 204	ZHENGXIAO WU	6
## 205	YU QIU LIU	6
## 206	YIFAN MEI	6
## 207	XINWEI DENG	6
## 208	XIAOPING FENG	6
## 209	VICTORIA MANSFIELD	6
## 210	TUN LEE NG	6
## 211	STEPHEN BERG	6
## 212	STEPHEN AARON STANHOPE	6
## 213	SHULEI WANG	6
## 214	SEAN KENT	6
## 215	RUNGANG HAN	6
## 216	PHILIP WOOD	6
## 217	MARIA ROJO	6
## 218	LILI ZHENG	6
## 219	KRISHNAKUMAR BALASUBRAMANIAN	6
## 220	JENNIFER NGUYEN	6
## 221	COLE COOK	6
## 222	CHIEN-WEI CHEN	6
## 223	CHAN PARK	6
## 224	ANDREW LESLIE	6

## 225	ALYSSA DIGILIO	6
## 226	YONGJOON KIM	5
## 227	XU HE	5
## 228	TING YE	5
## 229	SHANE HUBLER	5
## 230	RUIFENG XU	5
## 231	QI JIANG	5
## 232	MICHAEL KUTZLER	5
## 233	JASON P FINE	5
## 234	DANIELE CAPPELLETTI	5
## 235	CHENSHENG KUANG	5
## 236	CHELSEY GREEN	5
## 237	ANRU ZHANG	5
## 238	ALEXANDER COVINGTON	5
## 239	ZHIGUO XIAO	4
## 240	XUEYAO CHEN	4
## 241	XINJIE HE	4
## 242	VARSHA KULKARNI	4
## 243	TONGHAI YANG	4
## 244	PAUL SAVARIAPPAN	4
## 245	MICHAEL RENE KOSOROK	4
## 246	LEI XU	4
## 247	KATHERINE GOODE	4
## 248	JIE WEI	4
## 249	IAN BRANSTAD RILEY	4
## 250	GENG LI	4
## 251	FANG FANG	4
## 252	COLIN LONGHURST	4
## 253	CLAIRE BOBST	4
## 254	BO HUANG	4
## 255	BIN ZHANG	4
## 256	BENJAMIN RECHT	4
## 257	ZHIGENG GENG	3
## 258	YUJIN CHUNG	3
## 259	YUCHANG WU	3
## 260	YALI WANG	3
## 261	YAJUAN SI	3
## 262	XUN ZHAO	3
## 263	WAI TONG FAN	3
## 264	TAERI UHM	3
## 265	RONGJUN ZHU	3
## 266	MITCHELL PAUKNER	3
## 267	MICHAEL LIOU	3
## 268	MANJUSHA KANCHARLA	3
## 269	LU MAO	3
## 270	LAN LUO	3
## 271	LAM HO	3
## 272	KYLE HERBET	3
## 273	KYLE HEBERT	3
## 274	KEVIN HASEGAWA ENG	3
## 275	KEN ONO	3
## 276	KAZUHIKO SHINKI	3
## 277	KAILEI CHEN	3
## 278	JUNGWON MUN	3

## 279	JOHN KANE	3
## 280	JEEA CHOI	3
## 281	GUANNAN SUN	3
## 282	GINA OH	3
## 283	DEREK NORTON	3
## 284	DEBRAJ DAS	3
## 285	DAVID DEMETS	3
## 286	CHIA-CHIEH LIN	3
## 287	CHAOQUN MEI	3
## 288	BI CHENG WU	3
## 289	BEN ADAM HAALAND	3
## 290	BEHZAD AALIPUR HAFSHEJANI	3
## 291	BEHZAD AALIPUR	3
## 292	AUGUST JENSEN	3
## 293	ALEXANDER KISELEV	3
## 294	ZIJIAN NI	2
## 295	YUNONG LIN	2
## 296	XIUFENG SHAO	2
## 297	XIAOWU DAI	2
## 298	WEI ZHENG	2
## 299	TING-LI LIN	2
## 300	STEPHEN WAINGER	2
## 301	STEFFEN LEMPP	2
## 302	SONA Z SWANSON	2
## 303	SHUYUN YE	2
## 304	SHUANG HUANG	2
## 305	SHAN LU	2
## 306	SCOTT JOSEPH HETZEL	2
## 307	ROBERT WAYNE GREEN	2
## 308	NORMAN DRAPER	2
## 309	NICHOLAS HENDERSON	2
## 310	MIN JUNG LEE	2
## 311	MIHAELA IFRIM	2
## 312	MICHAEL GEORGE ILTIS	2
## 313	MARIAN R FISHER	2
## 314	MARIA KAMENETSKY	2
## 315	LIN QI	2
## 316	LANCINE KONATE	2
## 317	KRISTEN CYFFKA	2
## 318	JUNHEE HAN	2
## 319	JOSEP GINEBRA	2
## 320	JEE YEON KIM	2
## 321	JASON RICHARD SWANSON	2
## 322	JAMES ANDERSON	2
## 323	HUI WANG	2
## 324	HUAIBAO FENG	2
## 325	GREGORIO MORENO-FLORES	2
## 326	GARY HOWARD SCHROEDER	2
## 327	FANGFANG WANG	2
## 328	EDWARD ERKER	2
## 329	CHRISTOPHER WAGNER	2
## 330	CHENGNING ZHANG	2
## 331	ARNOLD MILLER	2
## 332	ANNE BRUCKNER	2

## 333	YANNAN QIU	1
## 334	YANBING ZHENG	1
## 335	XIAODAN WEI	1
## 336	THU LE	1
## 337	SUKHENDU MEHROTRA	1
## 338	STEVEN SAM	1
## 339	SRIKANTHMADHAVAN ARAVAMUTHAN	1
## 340	SHUAI CHEN	1
## 341	SHI JIN	1
## 342	SCOTT HOTTOVY	1
## 343	SAMUEL STECHMANN	1
## 344	RUIFANG SONG	1
## 345	ROBERT R MEYER	1
## 346	ROBERT HARRON	1
## 347	RICHARD A BRUALDI	1
## 348	REBECCA POST	1
## 349	NELLIE LAUGHLIN	1
## 350	NATALIA DE LEON GATTI	1
## 351	MICHELLE HARRIS	1
## 352	MICHAEL HOGAN	1
## 353	MATTHEW BALLARD	1
## 354	MARI PALTA	1
## 355	LEV BORISOV	1
## 356	KEEGAN KORTHAUER	1
## 357	JUNHO LEE	1
## 358	JONATHON PETERSON	1
## 359	JOHN WILTSHIRE-GORDON	1
## 360	JIAN WU	1
## 361	JESSE THOMAS HOLZER	1
## 362	JAMES D KUELBS	1
## 363	HYUNSEUNG KANG	1
## 364	HUILIN HU	1
## 365	GUILHERME ROSA	1
## 366	GUANHUA CHEN	1
## 367	FLORIAN BERTRAND	1
## 368	ELOISA D CHAVAS	1
## 369	DONGHYUN LEE	1
## 370	DIETRICH UHLENBROCK	1
## 371	CHRISTINE SORKNESS	1
## 372	ANDREJ ZLATOS	1
## 373	ANATOLE BECK	1
## 374	ALEXANDER FISH	1
## 375	ALBRECHT KLEMM	1
## 376	ALBERTO DEL PIA	1
## 377	ALAN HUANG	1

```
nrow(professores)
```

```
## [1] 377
```

```
# SELECT DISTINCT - Seleciona apenas nomes únicos de professores, evitando duplicatas
# i.name as professor_nome - Pega o nome do professor da tabela instructors e renomeia a coluna
# COUNT(DISTINCT s.uuid) - Conta quantas seções/turmas únicas de STAT cada professor lecionou
# FROM instructors i - Começa pela tabela de professores (alias 'i' para facilitar)
# JOIN teachings - Relaciona professores com as aulas que ministraram
```



```

# JOIN sections - Conecta com as seções/turmas específicas
# JOIN course_offerings - Liga com as ofertas de curso
# JOIN subject_memberships - Relaciona ofertas de curso com as disciplinas
# JOIN subjects - Acessa a tabela de disciplinas para ver a abreviação
# WHERE sub.abbreviation = 'STAT' - Filtra apenas disciplinas de Estatística
# GROUP BY i.name - Agrupa resultados por professor para fazer a contagem
# ORDER BY total_disciplinas_stat DESC - Ordena do professor com mais STAT para o com menos
# dbGetQuery(conn, sql) - Executa a consulta SQL no banco de dados conectado

```

3. O GPA americano é definido numa escala de 0 a 4, em que A = 4, AB = 3.5, B = 3, BC = 2.5, C = 2, D = 1 e F = 0. Determinando a nota média de cada oferecimento pela ponderação da quantidade de alunos em cada extrato com os valores numéricos de cada conceito, indique (no que se referente a disciplinas no assunto de estatística):

```

# --- 1) Professores: calcular média de notas (GPA) por professor e retornar todos os empates min/max
sql_professores <- "
WITH estat_prof AS (
  SELECT
    i.id AS professor_id,
    i.name AS professor_nome,
    SUM( (gd.a_count*4.0) + (gd.ab_count*3.5) + (gd.b_count*3.0) + (gd.bc_count*2.5) +
          (gd.c_count*2.0) + (gd.d_count*1.0) + (gd.f_count*0.0) ) AS total_pontos,
    SUM( (gd.a_count) + (gd.ab_count) + (gd.b_count) + (gd.bc_count) +
          (gd.c_count) + (gd.d_count) + (gd.f_count) ) AS total_alunos
  FROM subjects sub
  JOIN subject_memberships sm ON sm.subject_code = sub.code
  JOIN course_offerings co ON co.uuid = sm.course_offering_uuid
  JOIN sections s ON s.course_offering_uuid = co.uuid
  JOIN teachings t ON t.section_uuid = s.uuid
  JOIN instructors i ON i.id = t.instructor_id
  JOIN grade_distributions gd ON gd.course_offering_uuid = co.uuid AND gd.section_number = s.number
  WHERE sub.abbreviation = 'STAT'
  GROUP BY i.id
),
prof_media AS (
  SELECT
    professor_id,
    professor_nome,
    total_alunos,
    total_pontos,
    CASE WHEN total_alunos > 0 THEN total_pontos * 1.0 / total_alunos ELSE NULL END AS media_gpa
  FROM estat_prof
  WHERE total_alunos > 0
)
SELECT * FROM prof_media
ORDER BY media_gpa ASC;
"

professores_todos <- dbGetQuery(conn, sql_professores)
if (nrow(professores_todos) > 0) {
  gpa_min <- min(professores_todos$media_gpa, na.rm = TRUE)
  gpa_max <- max(professores_todos$media_gpa, na.rm = TRUE)
  professores_dificeis <- subset(professores_todos, media_gpa == gpa_min) # mais difíceis
  professores_faceis <- subset(professores_todos, media_gpa == gpa_max) # mais fáceis
}

```

```

} else {
  professores_difíceis <- professores_todos
  professores_fáceis <- professores_todos
}

# --- 2) Disciplinas: calcular média de notas (GPA) por disciplina e retornar empates
sql_disciplinas <- "
WITH ofertas_estat AS (
  SELECT
    co.uuid AS oferta_id,
    co.course_uuid,
    SUM( (gd.a_count*4.0) + (gd.ab_count*3.5) + (gd.b_count*3.0) + (gd.bc_count*2.5) +
      (gd.c_count*2.0) + (gd.d_count*1.0) + (gd.f_count*0.0) ) AS total_pontos,
    SUM( (gd.a_count) + (gd.ab_count) + (gd.b_count) + (gd.bc_count) +
      (gd.c_count) + (gd.d_count) + (gd.f_count) ) AS total_alunos
  FROM course_offerings co
  JOIN subject_memberships sm ON sm.course_offering_uuid = co.uuid
  JOIN subjects sub ON sub.code = sm.subject_code
  JOIN grade_distributions gd ON gd.course_offering_uuid = co.uuid
  WHERE sub.abbreviation = 'STAT'
  GROUP BY co.uuid
),
disc_agg AS (
  SELECT
    co.course_uuid,
    c.name AS disciplina_nome,
    SUM(oe.total_pontos) AS soma_pontos,
    SUM(oe.total_alunos) AS soma_alunos,
    CASE WHEN SUM(oe.total_alunos) > 0 THEN SUM(oe.total_pontos) * 1.0 / SUM(oe.total_alunos) ELSE NULL
  FROM ofertas_estat oe
  JOIN course_offerings co ON co.uuid = oe.oferta_id
  JOIN courses c ON c.uuid = co.course_uuid
  GROUP BY co.course_uuid
)
SELECT * FROM disc_agg
ORDER BY media_gpa ASC;
"

disciplinas_todas <- dbGetQuery(conn, sql_disciplinas)
if (nrow(disciplinas_todas) > 0) {
  gpa_min_disc <- min(disciplinas_todas$media_gpa, na.rm = TRUE)
  gpa_max_disc <- max(disciplinas_todas$media_gpa, na.rm = TRUE)
  disciplinas_difíceis <- subset(disciplinas_todas, media_gpa == gpa_min_disc) # mais difíceis
  disciplinas_fáceis <- subset(disciplinas_todas, media_gpa == gpa_max_disc) # mais fáceis
} else {
  disciplinas_difíceis <- disciplinas_todas
  disciplinas_fáceis <- disciplinas_todas
}

```

• Quem é o professor mais difícil?

```
professores_difíceis
```

```
## professor_id professor_nome total_alunos total_pontos media_gpa
## 1 2601633 JAMES D KUELBS 40 105.5 2.6375
```

- Quem é o professor mais fácil?

```
professores_faceis
```

```
##      professor_id      professor_nome total_alunos total_pontos
## 205      3884290 SRIKANTHMADHAVAN ARAVAMUTHAN          16          64
## 206      4272783          GUANHUA CHEN          16          64
## 207      4426932          YAJUAN SI           11          44
##      media_gpa
## 205          4
## 206          4
## 207          4
```

- Qual é a disciplina mais difícil?

```
disciplinas_dificeis
```

```
##      course_uuid
## 8 3ef3ba60-426e-39ad-a61f-17d381b193d6
##      disciplina_nome soma_pontos soma_alunos media_gpa
## 8 Introduction to the Theory of Probability      6439.5      2231  2.886374
```

- Qual é a disciplina mais fácil?

```
disciplinas_faceis
```

```
##      course_uuid
## 74 126a81bd-34e1-35f8-82f4-5edb70c3ae8b
## 75 61e7c0a0-e62b-37fc-bed5-db70c90c4f92
## 76 74fb97c9-9e98-3e8a-b0d9-66123b5a8e33
## 77 99d2f325-3f56-397f-8c71-d4a2490a8e03
##      disciplina_nome soma_pontos
## 74 Nonparametric Statistics and Machine Learning Methods      96
## 75      Sample Survey Theory and Method      216
## 76      Data Science Practicum      48
## 77      Empir Proc&Semiparmtrc Infernc      32
##      soma_alunos media_gpa
## 74          24          4
## 75          54          4
## 76          12          4
## 77           8          4
```

4. Desconecte do banco de dados.

```
dbDisconnect(conn)
```