GGplot notes

Key info

Needs a data frame, variables, what type of visualization you want

Syntax

Ggplot (data frame to graph, define the axis and what data goes into each axis)

Ex

ggplot(data = gapminder,

mapping = aes(x = year,

y = lifeExp,

by=country)) +

geom\_point() +

geom\_line(mapping=aes(color=continent))

data gives the file you are working from

mapping = aes gives the x and y variables and also lets you subset data via a second variable

gem\_point() allows you to put points onto a graph

gem\_line() turns the graph into a line graph

mapping aes(color=continent) color codes the lines

scale\_x\_log10() this provides a scaling of the axis

geom\_smooth() this allows you to fit a line

can do linear, exponential, etc

geom\_smooth(method = ‘modelname’)

color = lets you color the values by a type of factor

facet\_wrap(~ year) allows you to graph a lot of stuff together

the ~ indicates break down the graphs by the all the levels of the factor after the tilde.

For example ~ year gives you a graph for each year

Looping through data

For loops

for (my\_var in 1:10) {

print(my\_var)

}

Paste allows you to combine together two types of data

Collapses everything into a single string no longer a vector

paste("Iams", 5)

[1] "Iams 5"

> for (i in 1:3){

+ for (j in letters[1:3])

+ print(paste(i,j))

+ }

[1] "1 a"

[1] "1 b"

[1] "1 c"

[1] "2 a"

[1] "2 b"

[1] "2 c"

[1] "3 a"

[1] "3 b"

[1] "3 c"

> for (cont in unique(gapminder$continent)){

+ min\_life <- min(gapminder[gapminder$continnet == cont, 'lifeExp'])

+ max\_life <- max(gapminder[gapminder$continent == cont, 'lifeExp'])

+ print(paste(cont, "has a life expectancy betwen", min\_life, "and", max\_life))

+ }

[1] "Asia has a life expectancy betwen Inf and 82.603"

[1] "Europe has a life expectancy betwen Inf and 81.757"

[1] "Africa has a life expectancy betwen Inf and 76.442"

[1] "Americas has a life expectancy betwen Inf and 80.653"

[1] "Oceania has a life expectancy betwen Inf and 81.235"

for (cont in unique(gapminder$continent)){

This fills the list of what ot iterate through with the unique names of the continent

min\_life <- min(gapminder[gapminder$continnet == cont, 'lifeExp'])

First iteration you look for only the life expectanxy values that lne up with ‘lifeExp’

Second you move over to Europe and look thorugh life expectancy

Conditional statements

if(x>0){

print("x is positive")

}

if(x>0){

print("x is positive")

} else{

print("x is negative")

}

if(x>0){

print("x is positive")

} else if (x<0) {

print("x is negative")

} else {

print("other")

}

R will not continue down the list if one condition is met. You have to be the most specific in the beginning of the loop. You can nest if statements in order to avoid the program from exiting after the first TRUE response

Nested if inside a for

for (i in c(-1,3,0,-7)){

if (i>0){

print("positive number")

} else if (i<0){

print("negative number")

} else {

print("zero")

}

}

Look for mean expectancies greater than average

You can subset outputs from summary using []

Ex

> summary(gapminder$gdpPercap)

Min. 1st Qu. Median Mean 3rd Qu. Max.

241.2 1202.1 3531.8 7215.3 9325.5 113523.1

> summary(gapminder$gdpPercap)[1]

Min.

241.1659

Tidyverse/dplyr

Glimpse() this lets you look at the type of column and top rows of info

> glimpse(gapminder)

Observations: 1,704

Variables: 6

$ country <fct> Afghanistan, Afghanistan, Afghanista…

$ year <int> 1952, 1957, 1962, 1967, 1972, 1977, …

$ pop <dbl> 8425333, 9240934, 10267083, 11537966…

$ continent <fct> Asia, Asia, Asia, Asia, Asia, Asia, …

$ lifeExp <dbl> 28.801, 30.332, 31.997, 34.020, 36.0…

$ gdpPercap <dbl> 779.4453, 820.8530, 853.1007, 836.19…

Select

Chooses columns in a data frame

10 1997

11 2002

12 2007

13 1952

14 1957

15 1962

16 1967

17 1972

18 1977

19 1982

20 1987

21 1992

22 1997

23 2002

24 2007

25 1952

26 1957

27 1962

28 1967

29 1972

30 1977

31 1982

32 1987

33 1992

34 1997

35 2002

36 2007

37 1952

38 1957

39 1962

40 1967

41 1972

42 1977

43 1982

44 1987

45 1992

46 1997

47 2002

48 2007

49 1952

50 1957

51 1962

52 1967

53 1972

54 1977

55 1982

56 1987

57 1992

58 1997

59 2002

60 2007

61 1952

62 1957

63 1962

64 1967

65 1972

66 1977

67 1982

68 1987

69 1992

70 1997

71 2002

72 2007

73 1952

74 1957

75 1962

76 1967

77 1972

78 1977

79 1982

80 1987

81 1992

82 1997

83 2002

84 2007

85 1952

86 1957

87 1962

88 1967

89 1972

90 1977

91 1982

92 1987

93 1992

94 1997

95 2002

96 2007

97 1952

98 1957

99 1962

100 1967

101 1972

102 1977

103 1982

104 1987

105 1992

106 1997

107 2002

108 2007

109 1952

110 1957

111 1962

112 1967

113 1972

114 1977

115 1982

116 1987

117 1992

118 1997

119 2002

120 2007

121 1952

122 1957

123 1962

124 1967

125 1972

126 1977

127 1982

128 1987

129 1992

130 1997

131 2002

132 2007

133 1952

134 1957

135 1962

136 1967

137 1972

138 1977

139 1982

140 1987

141 1992

142 1997

143 2002

144 2007

145 1952

146 1957

147 1962

148 1967

149 1972

150 1977

151 1982

152 1987

153 1992

154 1997

155 2002

156 2007

157 1952

158 1957

159 1962

160 1967

161 1972

162 1977

163 1982

164 1987

165 1992

166 1997

167 2002

168 2007

169 1952

170 1957

171 1962

172 1967

173 1972

174 1977

175 1982

176 1987

177 1992

178 1997

179 2002

180 2007

181 1952

182 1957

183 1962

184 1967

185 1972

186 1977

187 1982

188 1987

189 1992

190 1997

191 2002

192 2007

193 1952

194 1957

195 1962

196 1967

197 1972

198 1977

199 1982

200 1987

201 1992

202 1997

203 2002

204 2007

205 1952

206 1957

207 1962

208 1967

209 1972

210 1977

211 1982

212 1987

213 1992

214 1997

215 2002

216 2007

217 1952

218 1957

219 1962

220 1967

221 1972

222 1977

223 1982

224 1987

225 1992

226 1997

227 2002

228 2007

229 1952

230 1957

231 1962

232 1967

233 1972

234 1977

235 1982

236 1987

237 1992

238 1997

239 2002

240 2007

241 1952

242 1957

243 1962

244 1967

245 1972

246 1977

247 1982

248 1987

249 1992

250 1997

251 2002

252 2007

253 1952

254 1957

255 1962

256 1967

257 1972

258 1977

259 1982

260 1987

261 1992

262 1997

263 2002

264 2007

265 1952

266 1957

267 1962

268 1967

269 1972

270 1977

271 1982

272 1987

273 1992

274 1997

275 2002

276 2007

277 1952

278 1957

279 1962

280 1967

281 1972

282 1977

283 1982

284 1987

285 1992

286 1997

287 2002

288 2007

289 1952

290 1957

291 1962

292 1967

293 1972

294 1977

295 1982

296 1987

297 1992

298 1997

299 2002

300 2007

301 1952

302 1957

303 1962

304 1967

305 1972

306 1977

307 1982

308 1987

309 1992

310 1997

311 2002

312 2007

313 1952

314 1957

315 1962

316 1967

317 1972

318 1977

319 1982

320 1987

321 1992

322 1997

323 2002

324 2007

325 1952

326 1957

327 1962

328 1967

329 1972

330 1977

331 1982

332 1987

333 1992

334 1997

335 2002

336 2007

337 1952

338 1957

339 1962

340 1967

341 1972

342 1977

343 1982

344 1987

345 1992

346 1997

347 2002

348 2007

349 1952

350 1957

351 1962

352 1967

353 1972

354 1977

355 1982

356 1987

357 1992

358 1997

359 2002

360 2007

361 1952

362 1957

363 1962

364 1967

365 1972

366 1977

367 1982

368 1987

369 1992

370 1997

371 2002

372 2007

373 1952

374 1957

375 1962

376 1967

377 1972

378 1977

379 1982

380 1987

381 1992

382 1997

383 2002

384 2007

385 1952

386 1957

387 1962

388 1967

389 1972

390 1977

391 1982

392 1987

393 1992

394 1997

395 2002

396 2007

397 1952

398 1957

399 1962

400 1967

401 1972

402 1977

403 1982

404 1987

405 1992

406 1997

407 2002

408 2007

409 1952

410 1957

411 1962

412 1967

413 1972

414 1977

415 1982

416 1987

417 1992

418 1997

419 2002

420 2007

421 1952

422 1957

423 1962

424 1967

425 1972

426 1977

427 1982

428 1987

429 1992

430 1997

431 2002

432 2007

433 1952

434 1957

435 1962

436 1967

437 1972

438 1977

439 1982

440 1987

441 1992

442 1997

443 2002

444 2007

445 1952

446 1957

447 1962

448 1967

449 1972

450 1977

451 1982

452 1987

453 1992

454 1997

455 2002

456 2007

457 1952

458 1957

459 1962

460 1967

461 1972

462 1977

463 1982

464 1987

465 1992

466 1997

467 2002

468 2007

469 1952

470 1957

471 1962

472 1967

473 1972

474 1977

475 1982

476 1987

477 1992

478 1997

479 2002

480 2007

481 1952

482 1957

483 1962

484 1967

485 1972

486 1977

487 1982

488 1987

489 1992

490 1997

491 2002

492 2007

493 1952

494 1957

495 1962

496 1967

497 1972

498 1977

499 1982

500 1987

501 1992

502 1997

503 2002

504 2007

505 1952

506 1957

507 1962

508 1967

509 1972

510 1977

511 1982

512 1987

513 1992

514 1997

515 2002

516 2007

517 1952

518 1957

519 1962

520 1967

521 1972

522 1977

523 1982

524 1987

525 1992

526 1997

527 2002

528 2007

529 1952

530 1957

531 1962

532 1967

533 1972

534 1977

535 1982

536 1987

537 1992

538 1997

539 2002

540 2007

541 1952

542 1957

543 1962

544 1967

545 1972

546 1977

547 1982

548 1987

549 1992

550 1997

551 2002

552 2007

553 1952

554 1957

555 1962

556 1967

557 1972

558 1977

559 1982

560 1987

561 1992

562 1997

563 2002

564 2007

565 1952

566 1957

567 1962

568 1967

569 1972

570 1977

571 1982

572 1987

573 1992

574 1997

575 2002

576 2007

577 1952

578 1957

579 1962

580 1967

581 1972

582 1977

583 1982

584 1987

585 1992

586 1997

587 2002

588 2007

589 1952

590 1957

591 1962

592 1967

593 1972

594 1977

595 1982

596 1987

597 1992

598 1997

599 2002

600 2007

601 1952

602 1957

603 1962

604 1967

605 1972

606 1977

607 1982

608 1987

609 1992

610 1997

611 2002

612 2007

613 1952

614 1957

615 1962

616 1967

617 1972

618 1977

619 1982

620 1987

621 1992

622 1997

623 2002

624 2007

625 1952

626 1957

627 1962

628 1967

629 1972

630 1977

631 1982

632 1987

633 1992

634 1997

635 2002

636 2007

637 1952

638 1957

639 1962

640 1967

641 1972

642 1977

643 1982

644 1987

645 1992

646 1997

647 2002

648 2007

649 1952

650 1957

651 1962

652 1967

653 1972

654 1977

655 1982

656 1987

657 1992

658 1997

659 2002

660 2007

661 1952

662 1957

663 1962

664 1967

665 1972

666 1977

667 1982

668 1987

669 1992

670 1997

671 2002

672 2007

673 1952

674 1957

675 1962

676 1967

677 1972

678 1977

679 1982

680 1987

681 1992

682 1997

683 2002

684 2007

685 1952

686 1957

687 1962

688 1967

689 1972

690 1977

691 1982

692 1987

693 1992

694 1997

695 2002

696 2007

697 1952

698 1957

699 1962

700 1967

701 1972

702 1977

703 1982

704 1987

705 1992

706 1997

707 2002

708 2007

709 1952

710 1957

711 1962

712 1967

713 1972

714 1977

715 1982

716 1987

717 1992

718 1997

719 2002

720 2007

721 1952

722 1957

723 1962

724 1967

725 1972

726 1977

727 1982

728 1987

729 1992

730 1997

731 2002

732 2007

733 1952

734 1957

735 1962

736 1967

737 1972

738 1977

739 1982

740 1987

741 1992

742 1997

743 2002

744 2007

745 1952

746 1957

747 1962

748 1967

749 1972

750 1977

751 1982

752 1987

753 1992

754 1997

755 2002

756 2007

757 1952

758 1957

759 1962

760 1967

761 1972

762 1977

763 1982

764 1987

765 1992

766 1997

767 2002

768 2007

769 1952

770 1957

771 1962

772 1967

773 1972

774 1977

775 1982

776 1987

777 1992

778 1997

779 2002

780 2007

781 1952

782 1957

783 1962

784 1967

785 1972

786 1977

787 1982

788 1987

789 1992

790 1997

791 2002

792 2007

793 1952

794 1957

795 1962

796 1967

797 1972

798 1977

799 1982

800 1987

801 1992

802 1997

803 2002

804 2007

805 1952

806 1957

807 1962

808 1967

809 1972

810 1977

811 1982

812 1987

813 1992

814 1997

815 2002

816 2007

817 1952

818 1957

819 1962

820 1967

821 1972

822 1977

823 1982

824 1987

825 1992

826 1997

827 2002

828 2007

829 1952

830 1957

831 1962

832 1967

833 1972

834 1977

835 1982

836 1987

837 1992

838 1997

839 2002

840 2007

841 1952

842 1957

843 1962

844 1967

845 1972

846 1977

847 1982

848 1987

849 1992

850 1997

851 2002

852 2007

853 1952

854 1957

855 1962

856 1967

857 1972

858 1977

859 1982

860 1987

861 1992

862 1997

863 2002

864 2007

865 1952

866 1957

867 1962

868 1967

869 1972

870 1977

871 1982

872 1987

873 1992

874 1997

875 2002

876 2007

877 1952

878 1957

879 1962

880 1967

881 1972

882 1977

883 1982

884 1987

885 1992

886 1997

887 2002

888 2007

889 1952

890 1957

891 1962

892 1967

893 1972

894 1977

895 1982

896 1987

897 1992

898 1997

899 2002

900 2007

901 1952

902 1957

903 1962

904 1967

905 1972

906 1977

907 1982

908 1987

909 1992

910 1997

911 2002

912 2007

913 1952

914 1957

915 1962

916 1967

917 1972

918 1977

919 1982

920 1987

921 1992

922 1997

923 2002

924 2007

925 1952

926 1957

927 1962

928 1967

929 1972

930 1977

931 1982

932 1987

933 1992

934 1997

935 2002

936 2007

937 1952

938 1957

939 1962

940 1967

941 1972

942 1977

943 1982

944 1987

945 1992

946 1997

947 2002

948 2007

949 1952

950 1957

951 1962

952 1967

953 1972

954 1977

955 1982

956 1987

957 1992

958 1997

959 2002

960 2007

961 1952

962 1957

963 1962

964 1967

965 1972

966 1977

967 1982

968 1987

969 1992

970 1997

971 2002

972 2007

973 1952

974 1957

975 1962

976 1967

977 1972

978 1977

979 1982

980 1987

981 1992

982 1997

983 2002

984 2007

985 1952

986 1957

987 1962

988 1967

989 1972

990 1977

991 1982

992 1987

993 1992

994 1997

995 2002

996 2007

997 1952

998 1957

999 1962

1000 1967

[ reached 'max' / getOption("max.print") -- omitted 704 rows ]

> head(select(gapminder, year))

year

1 1952

2 1957

3 1962

4 1967

5 1972

6 1977

You can add in as many columns as you want

Using a comma

> head(select(gapminder, year, lifeExp))

year lifeExp

1 1952 28.801

2 1957 30.332

3 1962 31.997

4 1967 34.020

5 1972 36.088

6 1977 38.438

You can remove a colm using –

> head(select(gapminder, -year))

country pop continent lifeExp gdpPercap

1 Afghanistan 8425333 Asia 28.801 779.4453

2 Afghanistan 9240934 Asia 30.332 820.8530

3 Afghanistan 10267083 Asia 31.997 853.1007

4 Afghanistan 11537966 Asia 34.020 836.1971

5 Afghanistan 13079460 Asia 36.088 739.9811

6 Afghanistan 14880372 Asia 38.438 786.1134

Filter

Allows you to select only one particular value of a factor

head(filter(gapminder, continent == "Asia"))

Or you can remove a factor value

> head(filter(gapminder, continent != "Asia"))

country year pop continent lifeExp gdpPercap

1 Albania 1952 1282697 Europe 55.23 1601.056

2 Albania 1957 1476505 Europe 59.28 1942.284

3 Albania 1962 1728137 Europe 64.82 2312.889

4 Albania 1967 1984060 Europe 66.22 2760.197

5 Albania 1972 2263554 Europe 67.69 3313.422

6 Albania 1977 2509048 Europe 68.93 3533.004

Or look for greater than

> head(filter(gapminder, year > 1997))

country year pop continent lifeExp gdpPercap

1 Afghanistan 2002 25268405 Asia 42.129 726.7341

2 Afghanistan 2007 31889923 Asia 43.828 974.5803

3 Albania 2002 3508512 Europe 75.651 4604.2117

4 Albania 2007 3600523 Europe 76.423 5937.0295

5 Algeria 2002 31287142 Africa 70.994 5288.0404

6 Algeria 2007 33333216 Africa 72.301 6223.3675

Pipeing

Mutate

Group by

Summarize