

Recoding Variables

```
> africa <- read.csv("data/africadata.csv", stringsAsFactors=FALSE)
> africa$region <- factor(africa$region)
> africa
```

	country	region	gdp_2017	pop_2017	area	rail	road
1	Algeria	north	99722535234	41063753	2381740	4691	113655.0
2	Angola	central	31374141168	26655513	1246700	NA	51429.0
3	Benin	west	4732594367	11458611	114760	758	19000.0
4	Botswana	south	11443275218	2343981	581730	888	25798.0
5	Burkina Faso	west	7013022439	19173322	274220	622	15271.6
6	Burundi	east	1215021410	11936481	27830	NA	12322.0
7	Cameroon	central	17940338306	24513689	475440	977	28857.0
8	Cape Verde	west	1129263677	533468	4030	NA	1350.0
9	Central African Republic	central	873967756	5098826	622980	NA	20278.0
10	Chad	central	4150723012	14965482	1284000	NA	40000.0
11	Comoros	east	294685870	825920	1861	NA	880.0
12	Congo, Dem. Rep.	central	10774647957	82242685	2344860	3641	153497.0
13	Congo, Rep.	central	5759941958	4866243	342000	795	17000.0
14	Cote d'Ivoire	west	18395731842	23815886	322460	639	81996.0
15	Djibouti	east	1194735681	911382	23200	781	3065.0
16	Egypt	north	194839401082	95215102	1001450	5195	137430.0

Recoding Variables

Recoding variables is a common task in data analysis, often used to transform categorical variables into numerical ones or to create new variables based on existing ones.

There are several methods for recoding variables, including using the `recode()` function in R, the `RECODE` command in SPSS, and the `PROC FREQ` statement in SAS.

Recoding variables can be useful for a variety of reasons, such as to simplify the data, to make it easier to interpret, or to create new variables that are more meaningful.

For example, you might recode a variable representing "gender" from "male" and "female" to "1" and "2" for use in a statistical model. Or you might recode a variable representing "age" from "18-24", "25-34", "35-44", "45-54", "55-64", and "65+" to "1", "2", "3", "4", "5", and "6" for use in a statistical model.