

# Análise da PNAD Contínua com o R

Aula ao Vivo 02 - 06 de Outubro de 2020

Vítor Wilher

Cientista de Dados | Mestre em Economia



# Plano de Voo

Introdução

Dados agregados da PNAD

Dados da PNAD: categoria do emprego

Dados da PNAD: Grupamento de atividades

Dados da PNAD: Rendimento

Dados da PNAD: Massa de Rendimentos

# Introdução

Nos dizeres do IBGE, a **PNAD Contínua** Visa acompanhar as flutuações trimestrais e a evolução, no curto, médio e longo prazos, da força de trabalho, e outras informações necessárias para o estudo do desenvolvimento socioeconômico do País. Para atender a tais objetivos, a pesquisa foi planejada para produzir indicadores trimestrais sobre a força de trabalho e indicadores anuais sobre temas suplementares permanentes (como trabalho e outras formas de trabalho, cuidados de pessoas e afazeres domésticos, tecnologia da informação e da comunicação etc.), investigados em um trimestre específico ou aplicados em uma parte da amostra a cada trimestre e acumulados para gerar resultados anuais, sendo produzidos, também, com periodicidade variável, indicadores sobre outros temas suplementares. Tem como unidade de investigação o domicílio.<sup>1</sup>

---

<sup>1</sup>Ver em site do IBGE.

# Introdução

*A PNAD Contínua foi implantada, experimentalmente, em outubro de 2011 e, a partir de janeiro de 2012, em caráter definitivo, em todo o Território Nacional. Sua amostra foi planejada de modo a produzir resultados para Brasil, Grandes Regiões, Unidades da Federação, Regiões Metropolitanas que contêm Municípios das Capitais, Região Integrada de Desenvolvimento - RIDE Grande Teresina, e Municípios das Capitais. Desde sua implantação, a pesquisa, gradualmente, vem ampliando os indicadores investigados e divulgados.*

# Introdução

Periodicidade de divulgação das informações:

- Mensal: Conjunto restrito de indicadores relacionados à força de trabalho e somente para o nível geográfico de Brasil;
- Trimestral: Conjunto de indicadores relacionados à força de trabalho para todos os níveis de divulgação da pesquisa;
- Anual: Demais temas permanentes da pesquisa e indicadores complementares à força de trabalho; e
- Variável: Outros temas ou tópicos dos temas permanentes a serem pesquisados com maior periodicidade ou ocasionalmente.

# Introdução

Os indicadores mensais utilizam as informações dos últimos três meses consecutivos da pesquisa, existindo, entre um trimestre móvel e o seguinte, repetição das informações de dois meses.

**Assim, os indicadores da PNAD Contínua produzidos mensalmente não refletem a situação de cada mês, mas, sim, a situação do trimestre móvel que finaliza a cada mês.**

# Dados agregados da PNAD

Os dados da PNAD podem ser acessadas através do site do SIDRA/IBGE. Aqui, vamos utilizar novamente o pacote `sidrar`.

```
## Pacotes utilizados nessa apresentação
library(tidyverse)
library(lubridate)
library(sidrar)
library(zoo)
library(scales)
library(timetk)
library(knitr)
```

# Dados agregados da PNAD

Para começar, nós pegamos os dados da **população total**:

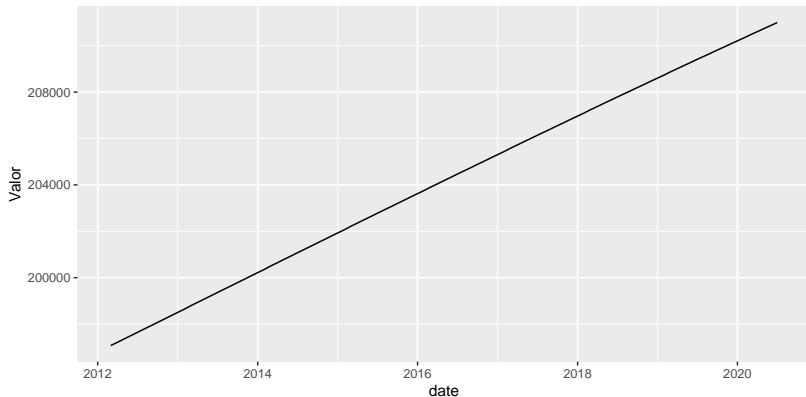
```
populacao = get_sidra(api='/t/6022/n1/all/v/606/p/all') %>%  
  mutate(date = parse_date(`Trimestre Móvel (Código)`,  
                             format='%Y%m')) %>%  
  select(date, Valor) %>%  
  as_tibble()  
tail(populacao)
```

```
## # A tibble: 6 x 2  
##   date      Valor  
##   <date>    <dbl>  
## 1 2020-02-01 210342  
## 2 2020-03-01 210474  
## 3 2020-04-01 210606  
## 4 2020-05-01 210738  
## 5 2020-06-01 210869  
## 6 2020-07-01 211001
```



# Dados agregados da PNAD

```
ggplot(populacao, aes(x=date, y=Valor))+  
  geom_line()
```



# Dados agregados da PNAD

Na sequência, vamos pegar os dados referentes à condição em relação à força de trabalho e condição de ocupação, disponíveis na tabela 6318.

```
names = c("date", 'pnea', 'pea', 'desocupada', 'ocupada', 'pia')
condicao = get_sidra(api='/t/6318/n1/all/v/1641/p/all/c629/all') %>%
  mutate(date = parse_date(`Trimestre Móvel (Código)`,
                           format='%Y%m')) %>%
  select(date, "Condição em relação à força de trabalho e condição de ocupação", Valor) %>%
  spread("Condição em relação à força de trabalho e condição de ocupação", Valor) %>%
  `colnames<-`(names) %>%
  as_tibble()
```

# Dados agregados da PNAD

```
tail(condicao)
```

```
## # A tibble: 6 x 6
##   date      pnea    pea desocupada ocupada    pia
##   <date>    <dbl> <dbl>      <dbl>   <dbl>  <dbl>
## 1 2020-02-01 65937 106052    12343   93710 171989
## 2 2020-03-01 67281 105073    12850   92223 172354
## 3 2020-04-01 70926 102052    12811   89241 172978
## 4 2020-05-01 74964  98646    12710   85936 173610
## 5 2020-06-01 77781  96138    12791   83347 173918
## 6 2020-07-01 78956  95158    13130   82027 174114
```

# Dados agregados da PNAD

Vamos agora organizar os dados:

```
agregado_pnad = inner_join(populacao, condicao, by='date') %>%  
  rename(populacao = Valor) %>%  
  mutate(inativos = populacao - pia,  
         desemprego = desocupada/pea*100,  
         participacao = pea/pia*100) %>%  
  select(date, populacao, inativos, pia, pea, pnea, ocupada, desocupada,  
         desemprego, participacao)
```

# Dados agregados da PNAD

```
agregado_pnad %>%  
  select(-desemprego, -participacao) %>%  
  tail() %>%  
  kable()
```

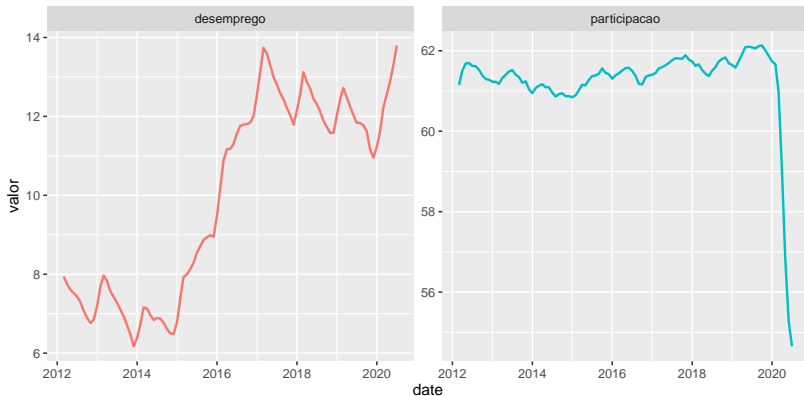
date	populacao	inativos	pia	pea	pnea	ocupada	desocupada
2020-02-01	210342	38353	171989	106052	65937	93710	12343
2020-03-01	210474	38120	172354	105073	67281	92223	12850
2020-04-01	210606	37628	172978	102052	70926	89241	12811
2020-05-01	210738	37128	173610	98646	74964	85936	12710
2020-06-01	210869	36951	173918	96138	77781	83347	12791
2020-07-01	211001	36887	174114	95158	78956	82027	13130

```
agregado_pnad %>%  
  select(date, desemprego, participacao) %>%  
  tail() %>%  
  kable()
```

date	desemprego	participacao
2020-02-01	11.63863	61.66208
2020-03-01	12.22959	60.96348
2020-04-01	12.55340	58.99710
2020-05-01	12.88446	56.82046
2020-06-01	13.30483	55.27777
2020-07-01	13.79810	54.65270

# Dados agregados da PNAD

```
agregado_pnad_long =  
  agregado_pnad %>%  
    gather(variavel, valor, -date)  
  
filter(agregado_pnad_long,  
       variavel %in% c('desemprego', 'participacao')) %>%  
ggplot(aes(x=date, y=valor, colour=variavel))+  
  geom_line(size=.8)+  
  facet_wrap(~variavel, scales='free')+  
  theme(legend.position = 'none')
```



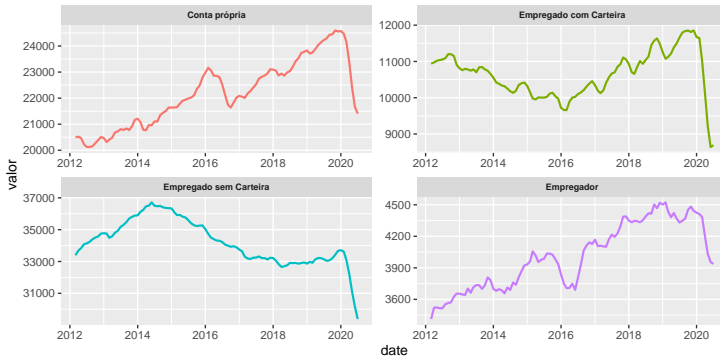
# Dados da PNAD: categoria do emprego

Agora, podemos pegar os dados referentes a **posição na ocupação e categoria do emprego no trabalho principal**, disponíveis na tabela 6320:

```
ocupacao_categoria =  
  get_sidra(api='/t/6320/n1/all/v/4090/p/all/c11913/allxt') %>%  
  mutate(date = parse_date(`Trimestre Móvel (Código)`,  
                             format='%Y%m')) %>%  
  select(date, "Posição na ocupação e categoria do emprego no trabalho principal", Valor) %>%  
  spread("Posição na ocupação e categoria do emprego no trabalho principal", Valor) %>%  
  rename('Empregado sem Carteira' = 'Empregado no setor privado, exclusive trabalhador doméstico - com carteira',  
         "Empregado com Carteira" = 'Empregado no setor privado, exclusive trabalhador doméstico - sem carteira')  
  as_tibble()
```

# Dados da PNAD: categoria do emprego

```
ocupacao_categoria_long =  
  ocupacao_categoria %>%  
    gather(variavel, valor, -date)  
  
filter(ocupacao_categoria_long,  
  variavel %in% c('Conta própria', 'Empregado sem Carteira',  
    "Empregado com Carteira",  
    "Empregador")) %>%  
  ggplot(aes(x=date, y=valor, colour=variavel))+  
  geom_line(size=.8)+  
  facet_wrap(~variavel, scales='free')+  
  theme(legend.position = 'none',  
    strip.text = element_text(size=7, face='bold'))
```





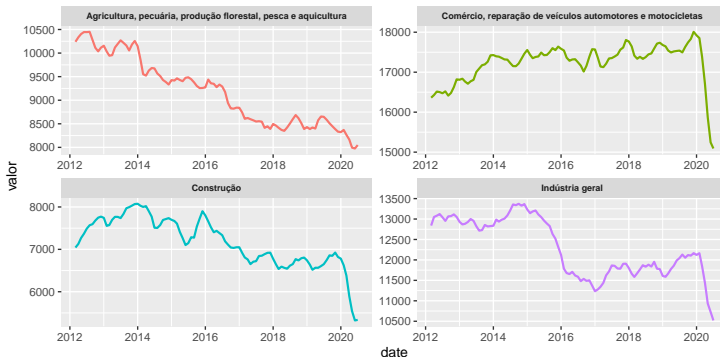
# Dados da PNAD: Grupamento de atividades

Agora, podemos pegar os dados referentes a **Grupamento de atividades no trabalho principal** , disponíveis na tabela 6323:

```
ocupacao_atividades =  
  get_sidra(api='/t/6323/n1/all/v/4090/p/all/c693/allxt') %>%  
  mutate(date = parse_date(`Trimestre Móvel (Código)`,  
    format='%Y%m')) %>%  
  select(date, "Grupamento de atividades no trabalho principal - PNADC",  
    Valor) %>%  
  spread("Grupamento de atividades no trabalho principal - PNADC",  
    Valor) %>%  
  as_tibble()
```

# Dados da PNAD: Grupamento de atividades

```
ocupacao_atividades_long =  
  ocupacao_atividades %>%  
    gather(variavel, valor, -date)  
  
filter(ocupacao_atividades_long,  
  variavel %in% c('Agricultura, pecuária, produção florestal, pesca e aquicultura', 'Comércio, reparação de veículos automotores e motocicletas', 'Indústria, construção e atividades de serviços', 'Serviços de saúde e educação', 'Serviços de recreação, cultura e lazer', 'Serviços de alojamento e alimentação', 'Serviços de transporte e comunicação', 'Serviços de administração e atividades de suporte', 'Serviços de saúde e educação', 'Serviços de recreação, cultura e lazer', 'Serviços de alojamento e alimentação', 'Serviços de transporte e comunicação', 'Serviços de administração e atividades de suporte'))  
ggplot(aes(x=date, y=valor, colour=variavel))+  
  geom_line(size=.8)+  
  facet_wrap(~variavel, scales='free')+  
  theme(legend.position = 'none',  
    strip.text = element_text(size=7, face='bold'))
```



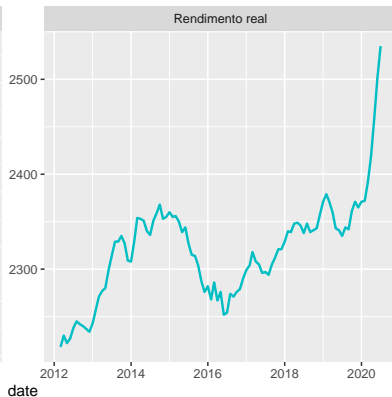
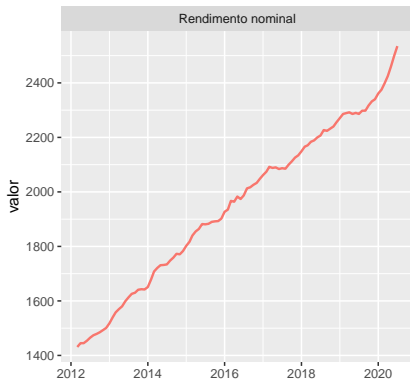
# Dados da PNAD: Rendimento

Agora, podemos pegar os dados referentes a **Rendimento médio, real e nominal, de todos os trabalhos, habitualmente recebido por mês**, disponíveis na tabela 6390:

```
pnad_rendimento =  
  get_sidra(api='/t/6390/n1/all/v/5929,5933/p/all') %>%  
  mutate(date = parse_date(`Trimestre Móvel (Código)`,  
                             format='%Y%m')) %>%  
  select(date, "Variável", Valor) %>%  
  spread("Variável", Valor) %>%  
  rename('Rendimento nominal' = "Rendimento médio nominal de todos os trabalhos, habitualmente recebido por  
        "Rendimento real" = "Rendimento médio real de todos os trabalhos, habitualmente recebido por mês  
  as_tibble()
```

# Dados da PNAD: Rendimento

```
pnad_rendimento_long =  
  pnad_rendimento %>%  
    gather(variavel, valor, -date)  
  
ggplot(pnad_rendimento_long,  
       aes(x=date, y=valor, colour=variavel))+  
  geom_line(size=.8)+  
  facet_wrap(~variavel, scales='free')+  
  theme(legend.position = 'none')
```



# Dados da PNAD: Massa de Rendimentos

Por fim, podemos pegar os dados referentes a **Massa de rendimento, real e nominal, de todos os trabalhos, habitualmente recebido por mês**, disponíveis na tabela 6392:

```
pnad_massa =  
  get_sidra(api='/t/6392/n1/all/v/6288,6293/p/all') %>%  
  mutate(date = parse_date(`Trimestre Móvel (Código)`,  
                             format='%Y%m')) %>%  
  select(date, "Variável", Valor) %>%  
  spread("Variável", Valor) %>%  
  rename('Massa de rendimento nominal' = "Massa de rendimento nominal de todos os trabalhos, habitualmente  
    "Massa de rendimento real" = "Massa de rendimento real de todos os trabalhos, habitualmente recebi  
  as_tibble()
```

# Dados da PNAD: Rendimento

```
pnad_massa_long =  
  pnad_massa %>%  
    gather(variavel, valor, -date)  
  
ggplot(pnad_massa_long,  
       aes(x=date, y=valor, colour=variavel))+  
  geom_line(size=.8)+  
  facet_wrap(~variavel, scales='free')+  
  theme(legend.position = 'none')
```

