

# Pesquisa de Survey

Fernando Meireles

📍 Sala Olavo Brasil

📅 Quartas-feiras, 9h-12h

🔗 [fmeireles.com/survey](https://fmeireles.com/survey)

## Apresentação

*Surveys* são, hoje, uma das mais difundidas ferramentas de pesquisa social. Por meio da aplicação de questionários estruturados a uma amostra de determinada população, governos estimam taxas de desemprego e de pobreza<sup>1</sup>; políticos definem estratégias de campanha<sup>2</sup>; institutos de pesquisa<sup>3</sup> identificam tendências de comportamento de diferentes grupos sociais; e empresas investigam a satisfação de seus clientes e testam ações de marketing. Apenas nas eleições de 2024, quase 15 mil pesquisas eleitorais foram conduzidas Brasil afora, cerca de 600 delas apenas no segundo turno.<sup>4</sup> Pesquisas de *survey* também passaram a ser conduzidas de novas formas, via telefone ou internet, ao mesmo tempo em que alimentaram diferentes polêmicas recentes, como a da subestimação da votação de candidaturas de extrema-direita<sup>5</sup> – ou, mais grave, as discussões sobre a proibição da divulgação de pesquisas eleitorais.<sup>6</sup> É difícil negar, enfim, que *surveys* são um pilar essencial da produção e difusão de conhecimento atualmente.

A proposta deste curso é oferecer uma introdução panorâmica a este tipo de pesquisa. Seu objetivo central é equipar alunos e alunas com conhecimentos básicos sobre desenho, implementação e análise pesquisas de *survey*. Entre os tópicos que veremos estão:

- **Total survey error**, isto é, diferentes tipos de erros que podem ocorrer em pesquisas de *survey* e seus diferentes *trade-offs*;
- **Amostragem** simples, estratificada, por conglomerados, em estágios e não-probabilística;
- **Mensuração e desenho de questionários**, incluindo elaboração de perguntas, *social desirability*, *nonresponse bias* e a validação de instrumentos;

---

<sup>1</sup>Vide a **PNADc**, por exemplo.

<sup>2</sup>Ver, por exemplo, o **Clube Associativo dos Profissionais de Marketing Político**

<sup>3</sup>A título de exemplo, a Associação Brasileira de Empresas de Pesquisa (ABEP) **tem atualmente 114 empresas listadas em atuação no Brasil.**

<sup>4</sup>**Metrópoles.**

<sup>5</sup>**Político.**

<sup>6</sup>**Senado.**

- **Ajustes e pós-estratificação** para ajustar uma amostra quando há fontes de erros conhecidas;
- **Tópicos específicos** como não-resposta e modelos de *likely voter*.

As aulas do curso integrarão teoria e tarefas aplicadas. Além de desenvolver habilidades específicas de pesquisa, alunos e alunas aplicarão o conteúdo estudado em projetos que cobrirão diferentes aspectos de uma pesquisa de *survey*.

## Pré-requisitos

São necessários alguns conhecimentos prévios de estatística e de programação em R para acompanhar este curso. Para estudantes regulares do IESP-UERJ, o *pré-requisito obrigatório é ter concluído Lego I*, mas, para um melhor aproveitamento da disciplina, é desejável ter cursado *Lego II* também. Para estudantes externos(as), é necessário ter comprovadamente conhecimentos sobre probabilidade e inferência estatística, modelos de regressão e uso de R para simulações, manipulação e visualização de dados.

## Leituras

Leremos diferentes textos ao longo deste curso, mas o seguinte livro servirá como a nossa principal base:

- Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2011). *Survey methodology*. John Wiley & Sons.

## Avaliação

Este curso terá dois tipos de avaliação: i) *realização de 4 tarefas*; e, ii), apresentação de um texto de leitura sugerida. Não haverá trabalho final.

## Tarefas (80 pontos)

Para praticar o conteúdo visto, teremos *4 tarefas* práticas que deverão ser entregues ao longo do curso, cada uma valendo *15 pontos na nota final*. Estas tarefas envolverão a aplicação de técnicas de pesquisa e análise de *survey* usando R. São elas:

- **Tarefa 1:** desenho de planos amostrais para sorteio de amostras aleatória simples e estratificadas;
- **Tarefa 2:** criação de um instrumento de pesquisa para *survey online*, incluindo a definição de variáveis e a redação de perguntas;

- **Tarefa 3:** análise de um banco de dados de *survey*;
- **Tarefa 4:** prática de pós-estratificação usando ponderação e *multilevel regression and post-stratification* (MrP).

A entrega das tarefas (exceto a segunda) deverá ser feita por meio de um modelo em **quarto** que será disponibilizado no início da disciplina. A avaliação levará em conta principalmente a capacidade de implementar o conhecimento que cobriremos e de justificar por escrito, com base na literatura, escolhas feitas.

## **Apresentação e participação (20 pontos)**

### **Apresentação de textos (10 pontos)**

Em cada aula, um ou mais alunos e alunas deverá apresentar brevemente os achados e recomendações de um dos textos de leitura sugerida. A apresentação deverá durar entre 10 e 15 minutos e deverá ser feita de forma estruturada, usando slides (no máximo, cinco). A apresentação será avaliada com base na capacidade de extrair recomendações e boas práticas para pesquisas de *survey*.

### **Participação em sala de aula (10 pontos)**

Alunos e alunas também terão suas avaliações compostas pela participação em sala de aula. Em particular, é esperado que, com base nas leituras obrigatórias e sugeridas, alunos e alunas contribuam com perguntas, comentários e discussões – inclusive sobre as apresentações dos e das colegas.

## **Nota final**

A nota final será calculada pela soma dos pontos obtidos em todas as tarefas e apresentações.

## **Políticas**

### **Política de Gênero**

Em aulas de metodologia, homens frequentemente monopolizam a participação. Para evitar isso, seguiremos dois protocolos neste curso: para intervir, é necessário estender a mão; quando mulheres falam, colegas não as interrompem.

### **Atendimento a Necessidades Especiais**

Se você, por qualquer razão que seja, tenha necessidades ou solicitações individuais, não hesite em procurar auxílio, tanto por **e-mail** quanto pessoalmente. Farei o melhor possível para tornar este curso um ambiente de aprendizado acessível e inclusivo.

## Plágios

Apresentar texto ou código de outra pessoa como se fosse próprio, inclusive em trabalhos finais de disciplina, viola **normativas da nossa universidade**, o que pode resultar em sanções. Embora a colaboração entre colegas seja estimulada neste curso, evite copiar códigos inteiros de outra pessoa – tarefas e exercícios serão checados automaticamente para identificar similaridade e plágio de fontes na internet.

## Uso de LLMs (ChatGPT)

O uso de LLMs (como o **ChatGPT**) é permitido neste curso, mas é importante usá-los apenas *para complementar o seu aprendizado em aula*. Em vez de apenas colar o enunciado de uma questão no chat e delegar totalmente a sua resolução (o que eu muito provavelmente descobrirei), tente primeiro resolvê-la por conta própria. Se a consulta a textos ou a mim não for suficiente ou possível, aí, sim, use LLM como uma espécie de tutor. Exemplo: “Estou tentando implementar bootstrap em R para calcular margens de erro para uma estimativa de survey cujo desenho amostral é estratificado. O código que fiz até agora está em anexo, mas retorna o seguinte erro: MENSAGEM DE ERRO. Você pode me explicar didaticamente qual é o problema?”. Importante notar que, mesmo assim, LLMs não são livres de erros e podem fornecer informações incorretas.

## Plano das Aulas

### Parte I – Desenho

#### Aula 1 – Apresentação do curso

*Leituras sugeridas:*

- *A Workers’ Inquiry (1880)*, Karl Marx
- *História das pesquisas de mercado no Brasil (projeto ABEP)*
- *Public Opinion Polling Basics: A Short Course Explaining How Public Opinion Polling Works* (Pew Research Center)

#### Aula 2 – O que são e para que servem pesquisas de *survey*

*Leituras obrigatórias:*

- Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2011). *Survey methodology*. John Wiley & Sons. Cap. 1 [*livro*]
- Zaller, J., & Feldman, S. (1992). A simple theory of the survey response: Answering questions versus revealing preferences. *American journal of political science*, 579–616.

### *Leituras sugeridas:*

- Hillygus, D. S. (2011). The Evolution of Election Polling in the United States. *Public Opinion Quarterly*, 75(5), 962–981.
- Converse, P. E. (2006). The nature of belief systems in mass publics (1964). *Critical review*, 18(1-3), 1–74.

### **Aula 3 – Erros em survey**

#### *Leituras obrigatórias:*

- Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2011). *Survey methodology*. John Wiley & Sons. Cap. 2 [livro]
- Berinsky, A. J. (2017). Measuring public opinion with surveys. *Annual review of political science*, 20(1), 309–329.

#### *Leituras sugeridas:*

- Groves, R. M., & Lyberg, L. (2010). Total survey error: Past, present, and future. *Public opinion quarterly*, 74(5), 849–879.
- Wolf, C., Joye, D., Smith, T. W., & Fu, Y. (2016). *The SAGE handbook of survey methodology*. Sage. Cap. 3 [livro]

### **Aula 4 – Amostragem**

#### *Leituras obrigatórias:*

- Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2011). *Survey methodology*. John Wiley & Sons. Caps. 3–4 [livro]
- Rossi, P. H., Wright, J. D., & Anderson, A. B. (1983). *Handbook of survey research*. Academic press. Cap. 2 [livro]

#### *Leituras sugeridas:*

- **Como avaliar a qualidade das pesquisas de opinião pública no Brasil?, Neale El-Dash (Polling-Data)**
- Squire, P. (1988). Why the 1936 Literary Digest poll failed. *Public Opinion Quarterly*, 52(1), 125–133.
- Lusinchi, D. (2012). “President” Landon and the 1936 Literary Digest Poll: were automobile and telephone owners to blame? *Social Science History*, 36(1), 23–54.

## Aula 5 – Amostras probabilísticas e não-probabilísticas

### *Leituras obrigatórias:*

- Wolf, C., Joye, D., Smith, T. W., & Fu, Y. (2016). *The SAGE handbook of survey methodology*. Sage. Cap. 22 [livro]
- Baker, R., Brick, J. M., Bates, N. A., Battaglia, M., Couper, M. P., Dever, J. A., Gile, K. J., & Tourangeau, R. (2013). Summary report of the AAPOR task force on non-probability sampling. *Journal of survey statistics and methodology*, 1(2), 90–143.

### *Leituras sugeridas:*

- Jerit, J., & Barabas, J. (2023). Are Nonprobability Surveys Fit for Purpose? *Public Opinion Quarterly*, 87(3), 816–840.
- Lupu, N., & Michelitch, K. (2018). Advances in survey methods for the developing world. *Annual Review of Political Science*, 21(1), 195–214.
- Lupu, N., Montalvo, J. D., Seligson, M. A., Zechmeister, E. J., & Zhirkov, K. (2022). Comparing Two Common Approaches to Within-Household Sampling: A Field Experiment in Costa Rica. *Unpublished manuscript*. Available at [https://noamlupu.com/FM\\_vs\\_LB.pdf](https://noamlupu.com/FM_vs_LB.pdf).

## Aula 6 – Desenho de questionários

### *Leituras obrigatórias:*

- Krosnick, J. A. (2018). Questionnaire Design. Em D. L. Vannette & J. A. Krosnick (Orgs.), *The Palgrave Handbook of Survey Research* (p. 439–455). Springer International Publishing.
- Wolf, C., Joye, D., Smith, T. W., & Fu, Y. (2016). *The SAGE handbook of survey methodology*. Sage. Cap. 16 [livro]

### *Leituras sugeridas:*

- Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2011). *Survey methodology*. John Wiley & Sons., Caps. 7–8 [livro]
- Schaeffer, N. C., & Dykema, J. (2011). Questions for surveys: current trends and future directions. *Public opinion quarterly*, 75(5), 909–961.
- Baker, A., & Renno, L. (2019). Nonpartisans as false negatives: the mismeasurement of party identification in public opinion surveys. *The Journal of Politics*, 81(3), 906–922.
- Luskin, R. C., & Bullock, J. G. (2011). “Don’t know” means “don’t know”: DK responses and the public’s level of political knowledge. *The Journal of Politics*, 73(2), 547–557.

## Parte II – Análise

### Aula 7 – Inferência com *surveys*

#### *Leituras obrigatórias:*

- Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2011). *Survey methodology*. John Wiley & Sons. Cap. 10 [*livro*]
- Lumley, T. (2010). *Complex surveys: a guide to analysis using R*. John Wiley & Sons. Caps. 1–2 [*livro*]

#### *Leituras sugeridas:*

- Shirani-Mehr, H., Rothschild, D., Goel, S., & Gelman, A. (2018). Disentangling bias and variance in election polls. *Journal of the American Statistical Association*, 113(522), 607–614.
- Valliant, R., Dever, J. A., Kreuter, F., Valliant, R., Dever, J. A., & Kreuter, F. (2018). Non-probability sampling. *Practical tools for designing and weighting survey samples*, 565–603. [*livro*]

### Aula 8 – Pós-estratificação

#### *Leituras obrigatórias:*

- Lumley, T. (2010). *Complex surveys: a guide to analysis using R*. John Wiley & Sons. Cap. 7 [*livro*]
- Wolf, C., Joye, D., Smith, T. W., & Fu, Y. (2016). *The SAGE handbook of survey methodology*. Sage. Cap. 31 [*livro*]

#### *Leituras sugeridas:*

- [Survey Weights in AmericasBarometer \(LAPOP\)](#)
- [How different weighting methods work \(Pew Research Center\)](#)

### Aula 9 – Pós-estratificação com modelos

#### *Leituras obrigatórias:*

- Wang, W., Rothschild, D., Goel, S., & Gelman, A. (2015). Forecasting elections with non-representative polls. *International Journal of Forecasting*, 31(3), 980–991.
- Hanretty, C. (2020). An Introduction to Multilevel Regression and Post-Stratification for Estimating Constituency Opinion: *Political Studies Review*, 18(4), 630–645.

*Leituras sugeridas:*

- Lax, J. R., & Phillips, J. H. (2009). How Should We Estimate Public Opinion in the States. *American Journal of Political Science*, 53(1), 107–121.
- Buttice, M. K., & Highton, B. (2013). How does multilevel regression and poststratification perform with conventional national surveys? *Political analysis*, 21(04), 449–467.
- Lopez-Martin, J., Phillips, J. H., & Gelman, A. (2022). Multilevel regression and poststratification case studies. URL <https://bookdown.org/jl5522/MRP-case-studies/>, 902, 903.

## **Aula 10 – Modelos flexíveis de MrP**

*Leituras obrigatórias:*

- Ghitza, Y., & Gelman, A. (2013). Deep Interactions with MRP: Election Turnout and Voting Patterns Among Small Electoral Subgroups. *American Journal of Political Science*, 57(3), 762–776.
- Bisbee, J. (2019). BARP: Improving Mister P Using Bayesian Additive Regression Trees. *American Political Science Review*, 113(4), 1060–1065.

*Leituras sugeridas:*

- Goplerud, M. (2023). Re-Evaluating Machine Learning for MRP Given the Comparable Performance of (Deep) Hierarchical Models. *American Political Science Review*, 1–8.
- Leemann, L., & Wasserfallen, F. (2017). Extending the Use and Prediction Precision of Subnational Public Opinion Estimation. *American Journal of Political Science*, 61(4), 1003–1022.

## **Parte III – Tópicos em pesquisa de *survey***

### **Aula 11 – Não-resposta**

*Leituras obrigatórias:*

- Wolf, C., Joye, D., Smith, T. W., & Fu, Y. (2016). *The SAGE handbook of survey methodology*. Sage. Cap. 27 [livro]
- Curtin, R., Presser, S., & Singer, E. (2005). Changes in telephone survey nonresponse over the past quarter century. *Public opinion quarterly*, 69(1), 87–98.

*Leituras sugeridas:*

- Clinton, J. D., Lapinski, J. S., & Trussler, M. J. (2022). Reluctant Republicans, eager Democrats? Partisan nonresponse and the accuracy of 2020 presidential pre-election telephone polls. *Public Opinion Quarterly*, 86(2), 247–269.



- Cavari, A., & Freedman, G. (2023). Survey nonresponse and mass polarization: The consequences of declining contact and cooperation rates. *American Political Science Review*, 117(1), 332–339.
- Peytchev, A., Baxter, R. K., & Carley-Baxter, L. R. (2009). Not all survey effort is equal: Reduction of nonresponse bias and nonresponse error. *Public Opinion Quarterly*, 73(4), 785–806.

## **Aula 12 – Viés de desejabilidade social**

### *Leituras obrigatórias:*

- Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological bulletin*, 133(5), 859.
- Rosenfeld, B., Imai, K., & Shapiro, J. N. (2016). An empirical validation study of popular survey methodologies for sensitive questions. *American Journal of Political Science*, 60(3), 783–802.

### *Leituras sugeridas:*

- Coppock, A. (2017). Did shy Trump supporters bias the 2016 polls? Evidence from a nationally-representative list experiment. *Statistics, Politics and Policy*, 8(1), 29–40.
- Streb, M. J., Burrell, B., Frederick, B., & Genovese, M. A. (2008). Social desirability effects and support for a female American president. *Public Opinion Quarterly*, 72(1), 76–89.
- DeBell, M., Krosnick, J. A., Gera, K., Yeager, D. S., & McDonald, M. P. (2020). The turnout gap in surveys: Explanations and solutions. *Sociological Methods & Research*, 49(4), 1133–1162.
- Hadaway, C. K., Marler, P. L., & Chaves, M. (1993). What the polls don't show: A closer look at US church attendance. *American Sociological Review*, 741–752.
- **Measuring Gender Identity in Latin America and the Caribbean (LAPOP)**

## **Aula 13 – Pesquisas eleitorais, eleitorado provável e indecisos**

### *Leituras obrigatórias:*

- Kenett, R. S., Pfeffermann, D., & Steinberg, D. M. (2018). Election polls—a survey, a critique, and proposals. *Annual Review of Statistics and Its Application*, 5(1), 1–24.
- Rentsch, A., Schaffner, B. F., & Gross, J. H. (2019). The elusive likely voter: Improving electoral predictions with more informed vote-propensity models. *Public Opinion Quarterly*, 83(4), 782–804.

### *Leituras sugeridas:*

- Meireles, F., & Russo, G. (2022). Pesquisas eleitorais no Brasil: tendências e desempenho. *Estudos Avançados*, 36, 117–131.

- Pereira, F. B., & Nunes, F. (2024). Pesquisas eleitorais e mudanças tardias na decisão do voto. *Opinião Pública*, 30, e3011.