



INSTITUTO NACIONAL DE ESTATÍSTICA  
STATISTICS PORTUGAL

# Infraestrutura Nacional de Dados

## Desenvolvimentos recentes

# Agenda da apresentação

## O que é a IND?

- Conceito e objetivo.

## Desenvolvimentos Recentes

- Integração e tratamento de dados.
- Dados administrativos e investigação.
- Recursos envolvidos

## Próximos Passos

- Estratégias futuras.

## Conclusão

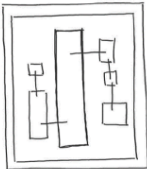
# Infraestrutura Nacional de Dados

A Infraestrutura Nacional de Dados aplica processos ETL (Extrair, Transformar, Carregar) para integrar, processar e unificar dados de múltiplas fontes, melhorando a análise e a tomada de decisões (lançada em 2018).

## Principais Objetivos:

- Aproveitar múltiplas fontes de dados.
- Aumentar a capacidade analítica e a flexibilidade na gestão.
- Integrar novas fontes para estatísticas e investigação.
- Incentivar a inovação.

**Visão Geral:** Reflete as estratégias do INE, criando um ecossistema de dados abrangente e integrado para atender a organismos públicos, empresas, investigadores e o público em geral.

Appropriation and use of administrative data and other sources	Partnerships	New roles
Data integration		Information security and data protection
Innovation	Integrate and share platforms, algorithms, knowledge, ...	Increase the economic and social value of the public good: statistical information

# Pilares da Infraestrutura Nacional de Dados

## Uso de Dados Administrativos e Outras Fontes

Aproveitamento de múltiplas fontes para ampliar a base de dados e enriquecer a análise.

## Parcerias

Fortalecimento de colaborações estratégicas para aumentar eficiência e promover inovação.

## Novos Papéis

Adaptação das funções para atender às novas procuras de dados e estatísticas.

## Integração de Dados

Unificação de fontes de dados para permitir análises mais robustas e precisas.

## Segurança e Proteção de Dados

Garantia de confidencialidade e segurança na manipulação e uso de informações.

## Inovação

Desenvolvimento de soluções inovadoras para maximizar o impacto dos dados.

## Integração e Partilha de Recursos

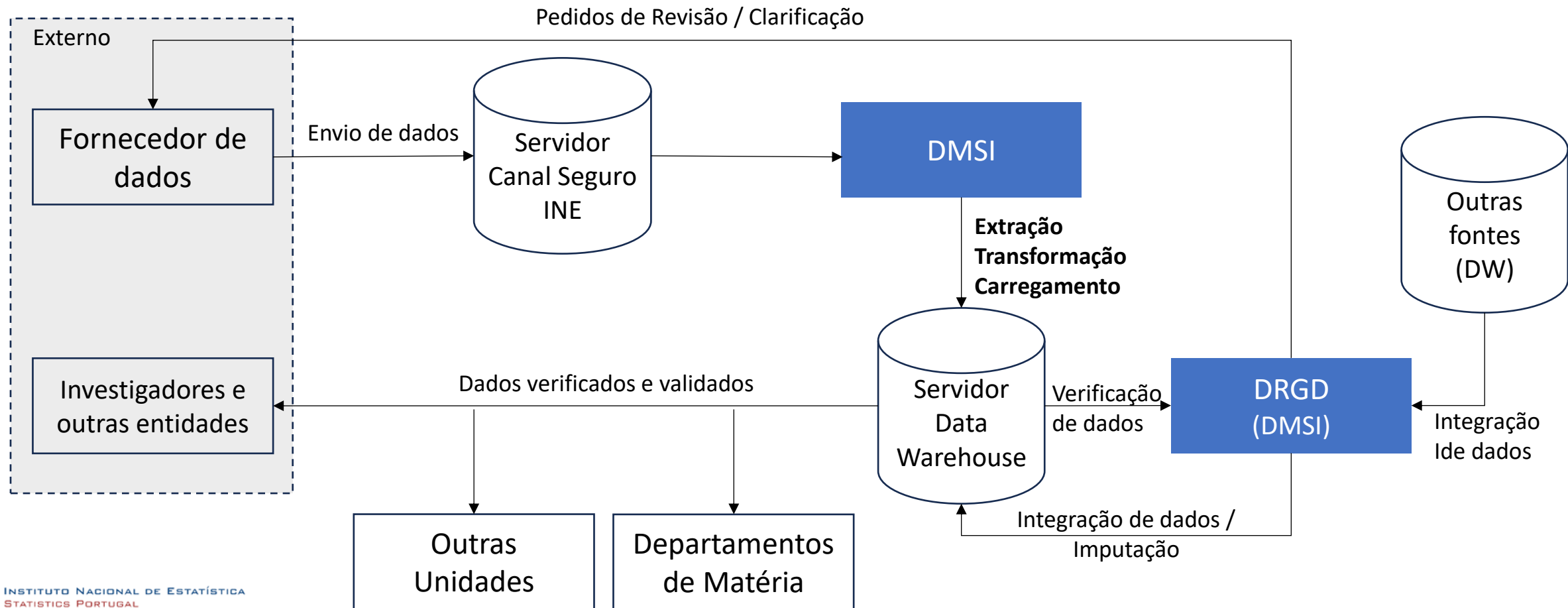
Plataformas, algoritmos e conhecimento acessíveis e colaborativos.

## Valorização do Bem Público

Aumento do valor económico e social das estatísticas oficiais para apoiar decisões públicas e privadas.

# Tratamento e Integração de fontes

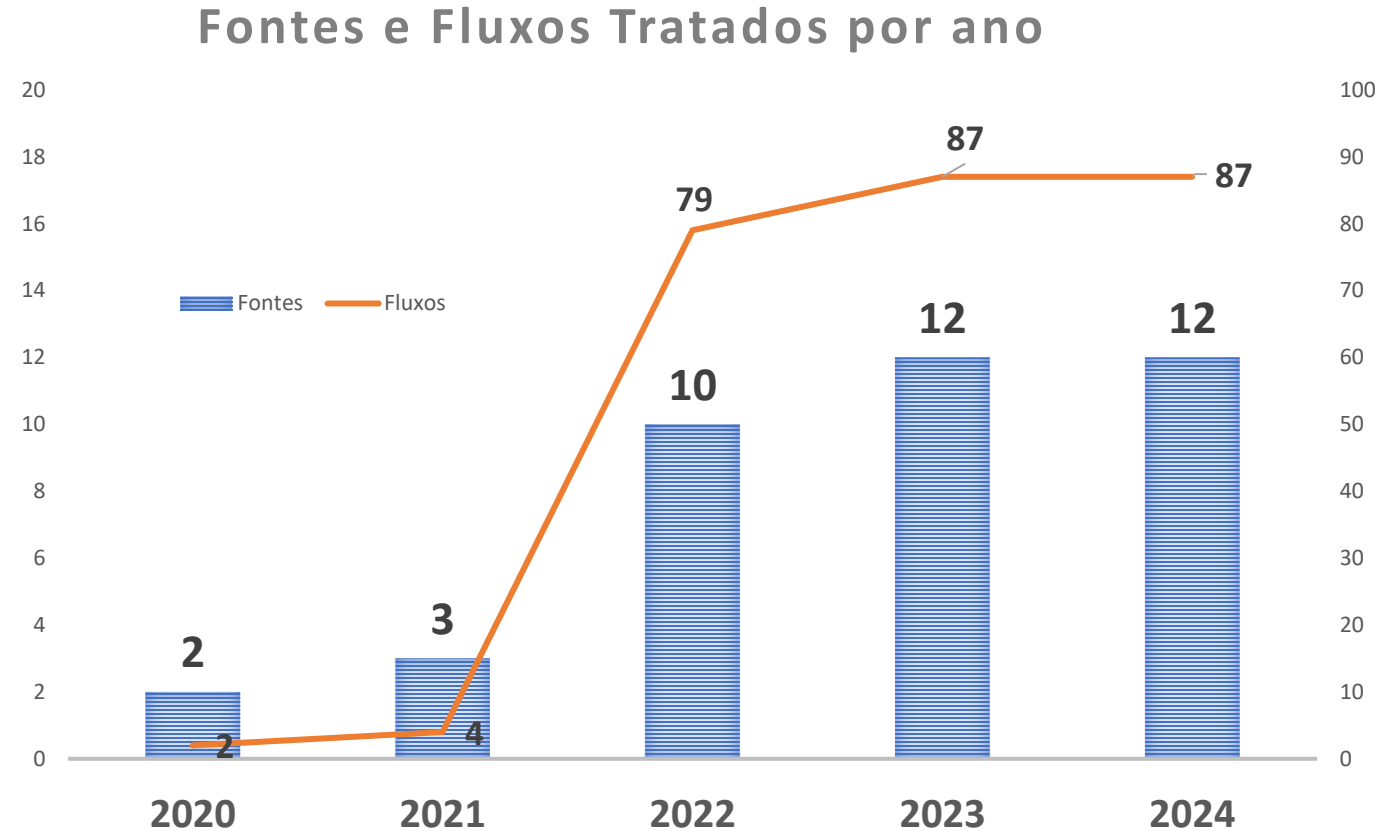
## Visão geral do processamento de Dados Administrativos



# Tratamento e Integração de fontes

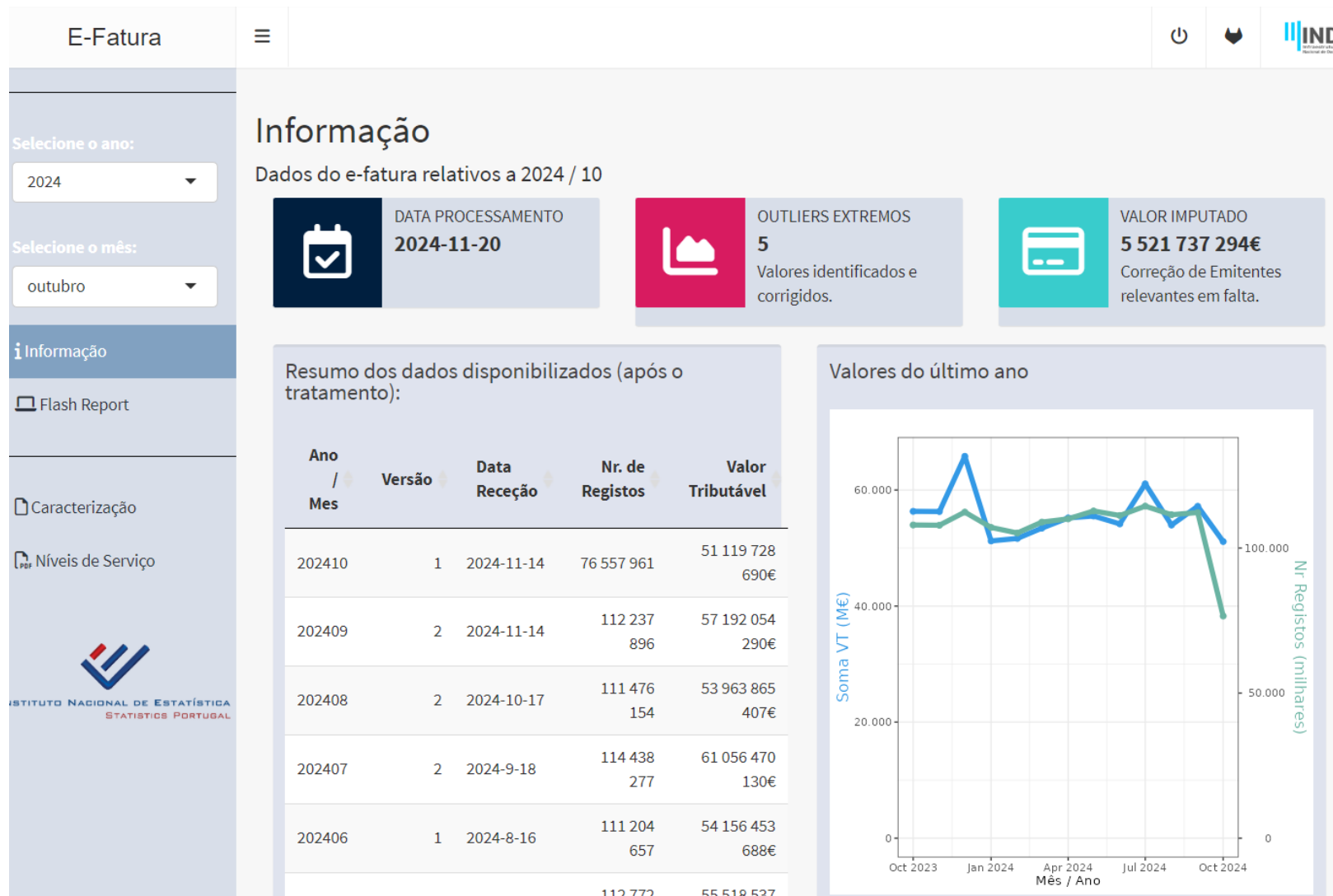
## Criação de Bases de Dados Consolidadas

- Dados aprimorados e acessíveis para todos os utilizadores (transversais).
  - Controlo de qualidade
  - Enriquecimento de variáveis
- Alinhamento com os procedimentos tradicionais na Operações Estatísticas:
  - Documento Metodológico
  - Níveis de Serviço
  - Flash Report (uso interno e externo)
  - Relatórios de Execução



# Tratamento e Integração de fontes

## Documentação dinâmica no exemplo do e-Fatura



# Desenvolvimentos recentes

## Integração de fontes administrativas nas Operações

Censos com Dados Administrativos

Base de População Residente

Base Nacional de Edifícios

Índice de Preços Habitação

Estatísticas de Preços da Habitação ao Nível Local

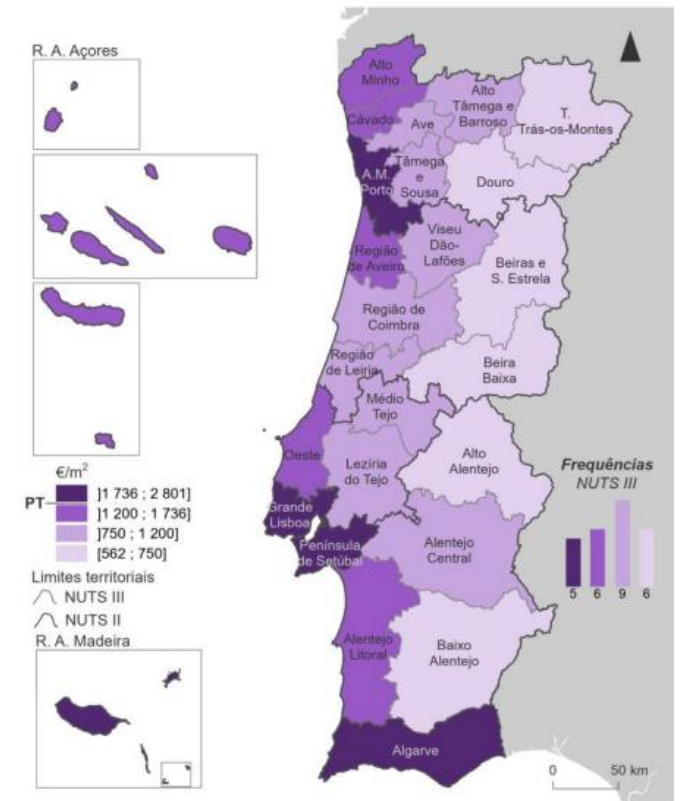
Remuneração bruta mensal média por trabalhador

(inicialmente publicado no STATSLAB em 2019, passou a estatística oficial em 2021)

Nascimento do primeiro filho: o que muda no padrão de despesas das famílias



Figura 2. Valor mediano das vendas por m<sup>2</sup> de alojamentos familiares em Portugal e NUTS III, 2.º trimestre de 2024

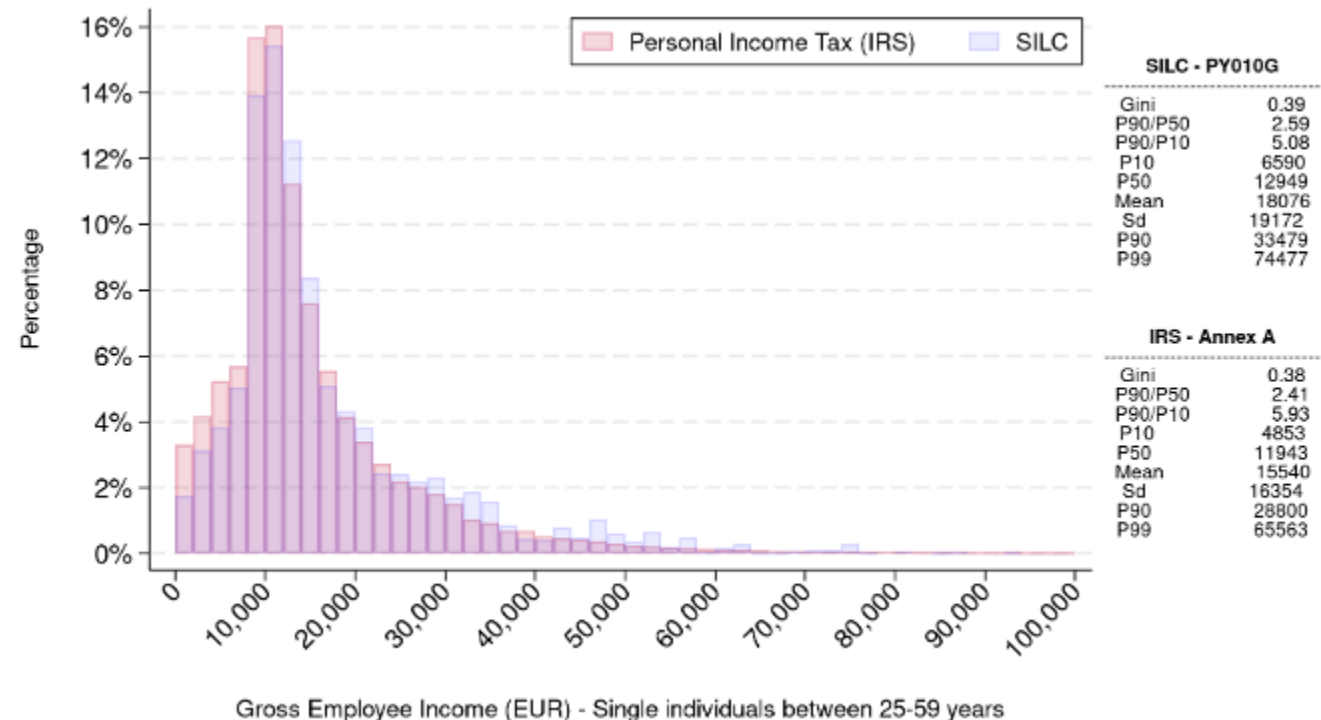
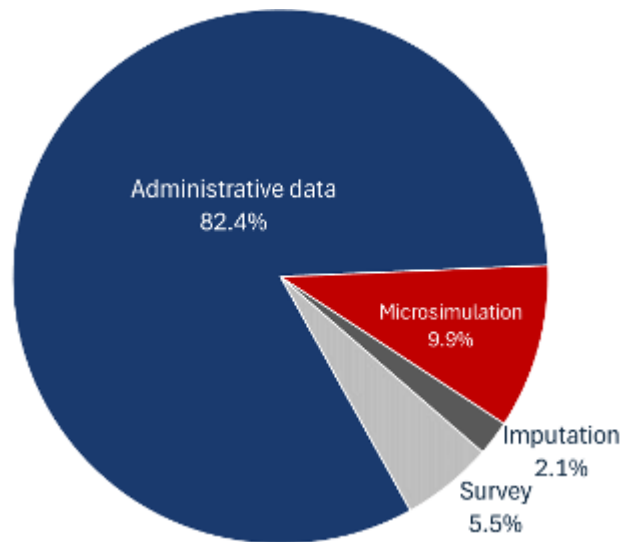




# Desenvolvimentos recentes

## Integração de fontes administrativas nas Operações

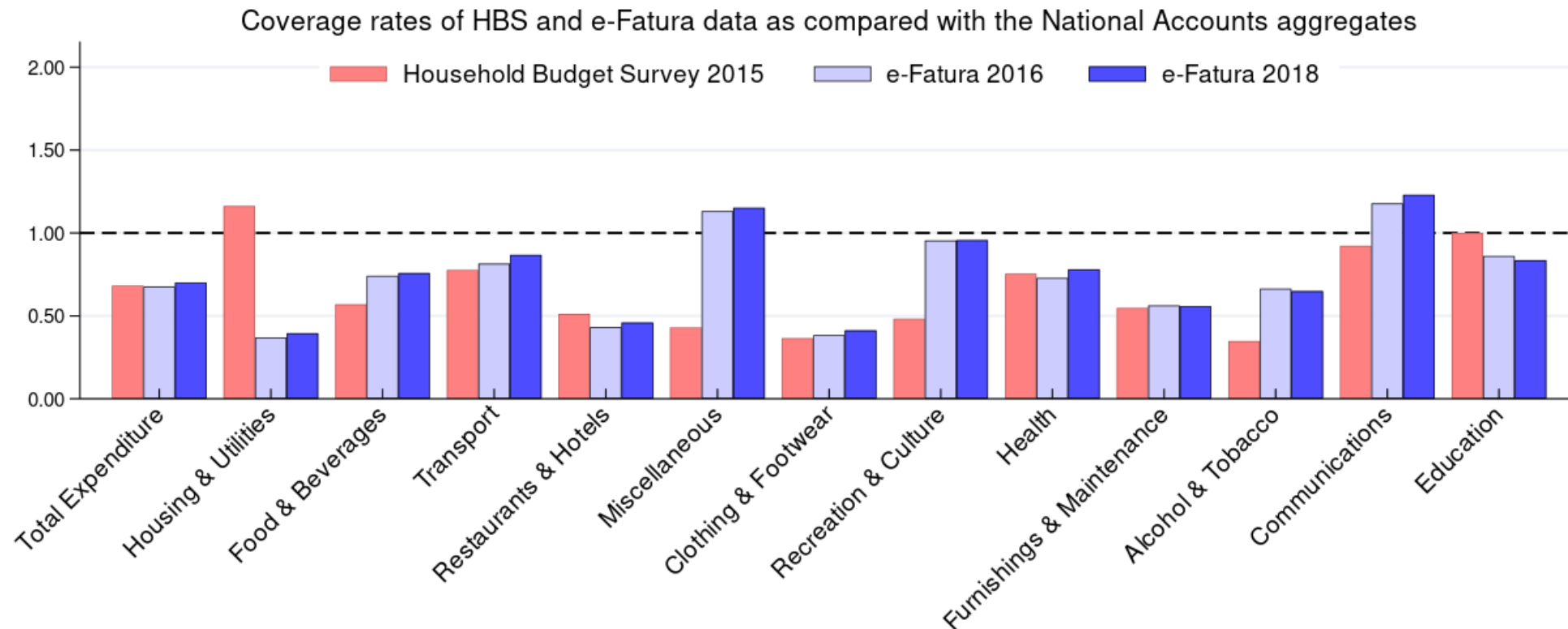
Integração de dados administrativos no Inquérito às Condições de vida e Rendimento (ICOR) – Rendimentos dos trabalhadores por conta de outrem. Em 2024, alargamento às pensões do regime contributivo.



# Desenvolvimentos recentes

## Integração de fontes administrativas nas Operações

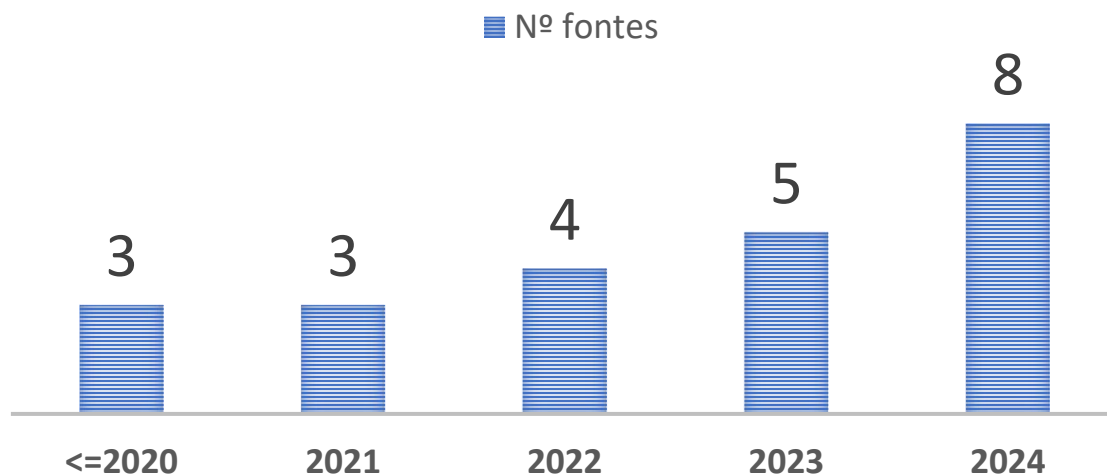
Integração de dados do e-fatura no  
Inquérito à Despesa dos Orçamentos Familiares (IDF) – em estudo



**Notes:** e-Fatura excludes actual and imputed rentals for housing. Sorted by product share in total expenditure.

# Disponibilização de dados à comunidade científica

## Fontes Administrativas Disponibilizadas



**Em 2022** IRS

**Em 2023** e-Fatura (emitente)

**Em 2024** Contratos de arrendamento  
Recibo Eletrónico de Rendas  
Comunicação Annual de Rendas Recebidas

## Exemplo: Ficha caracterização IRS – Anexo B

- 1 Objectivo
- 2 Informações Principais
- 3 Sumário Estatístico
- 4 Preenchimento dos Códigos CAE e CIRS (Quadro 3):

### Modelo 03 do IRS: Anexo B - 2022

DRGD / SDAE / NDA  
2023-09-19



#### 1 Objectivo

Descrever a informação disponibilizada para o Anexo B do Modelo 03 do IRS de 2022.

#### 2 Informações Principais

##### 2.1 Fonte, Versão, Unidade de Registo e Chave de Ligação

Impresso IRS Mod. 3 2022	Portaria 303/2021, de 17 de dezembro
Tabela	<a href="#">IRS_2022_ANEXO_B</a>
Versão	Setembro N+1
Unidade de registo das variáveis	Indivíduo (NIF_TITULAR)
Chave de ligação com o Rosto	<a href="#">HID NIF_TITULAR</a>

##### 2.2 Dicionário de Variáveis

###### 2.2.1 Descrição das Variáveis

Variavel	Tipo	Unidade	Descricao	N	Media
ANO	String	HID	B_Q02_01: ano dos rendimentos	1.301.374	
VERSÃO	String	HID	R_VARDERIV: variável	1.301.374	

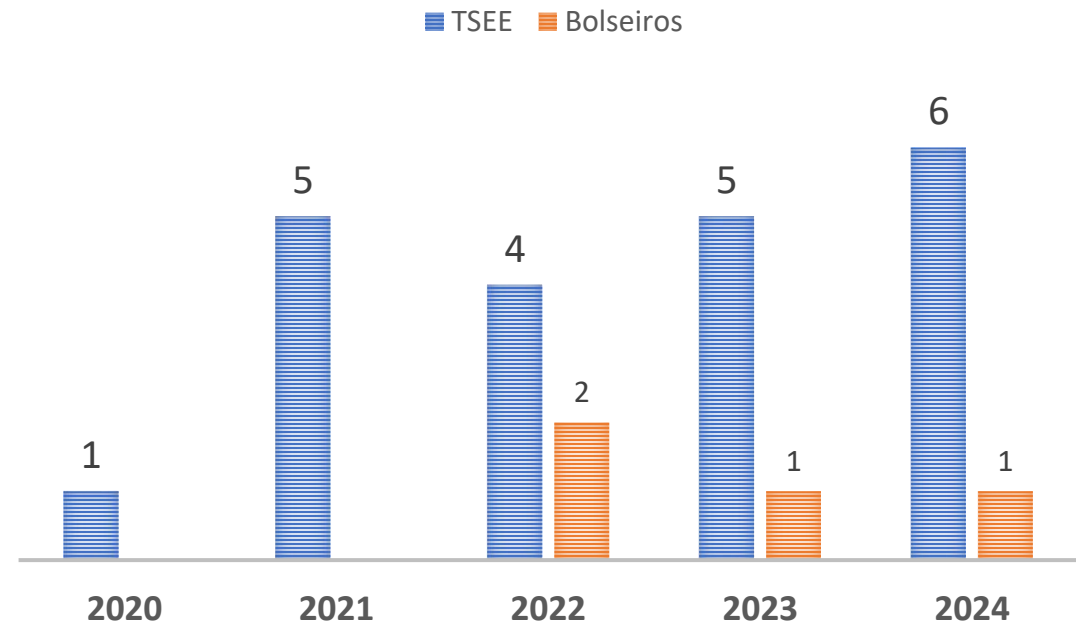
# Desenvolvimentos recentes

## Organização e Recursos Humanos

### Núcleo de Dados Administrativos (2020)

- Integrado no Departamento de Recolha e Gestão de Dados
- Identifica potenciais novas fontes de dados
- Define modelos de validação, consistência e coerência
- Procede à integração de dados de diferentes fontes
- Simplificação e padronização da utilização dos dados administrativos na produção de estatísticas

### Recursos Humanos Exclusivos



# Desenvolvimentos recentes

## Recursos Envolvidos

### Recrutamento especializado

- Enfoque nas competências de ciência de dados

### Suporte Adicional

- Investigadores residentes
- "Meia-hoRa" - reuniões semanais
  - Formação colaborativa em R

½ houR

half-an-hour of R

- Preface
- 1 Introduction
- 2 Code of conduct
- 3 Procedures
- 4 Calendar
- 5 Summary
- 6 Learning
- 7 ½ hoRa (2023)
- 8 References

Sponsored by:



INSTITUTO NACIONAL DE ESTATÍSTICA  
STATISTICS PORTUGAL



Partners:



**DGS** desde 1899  
Direção-Geral da Saúde



**INĒ**  
Instituto Nacional de Estatística



2024-12

December  
Theme: Special guests

Mo	Tu	We	Th	Fr	Sa	Su
2	3	4 Sandra Barragan SGM / INE	5	6	7	8
9	10	11 Rita Gaio FCUP	12	13	14	15
16	17	18 Kaija Burk Destatis	19	20	21	22
23	24	25	26	27	28	29
30	31					

cuRator: Mimoso Pinho; DMSI/GEO

# Desenvolvimentos recentes

## Formação Especializada

### Curso de admissão na carreira TSEE:

Duração de seis meses: teórico (2m) + prático (4m)

Aprendizagem das várias etapas de produção de estatísticas oficiais, participação na cadeia de processo, acompanhamento por um tutor

### Cursos especializados, em articulação direta com algumas das principais universidades



ISEC  
Lisbon School  
of Economics  
& Management  
UNIVERSIDADE DE LISBOA



Alguns exemplos de cursos especificamente desenhados para o INE:

**Em 2022** “Econometria”: 26 participantes, 30 horas

“Artificial Intelligence in Businesses” (1ª edição): 24 participantes, 35h

**Em 2023** “Macroeconometria”: 24 participantes, 30h

“Artificial Intelligence for Businesses” (2ª edition): 21 participantes, 35h

“Cibersegurança”: 20 participantes, 21h

**Em 2024** “Network Analytics and Deep Learning” : 24 participantes, 35h

# Próximos passos

- Tratamento de Dados:
  - Foco na sistematização e consistência das fontes ligadas à Construção e Habitação.
  - Aprofundamento da Base Nacional de Edifícios e dos Censos Administrativos.
- Integração de Novas Fontes:
  - Exemplo: Segurança Social.
- Continuidade na integração de dados administrativos nas operações estatísticas.

# Conclusão

- **Consolidação da IND**
  - Uma infraestrutura sólida, focada na integração e valorização de dados administrativos para uso estatístico e científico.
- **Impactos Positivos**
  - Apoio à decisão pública e privada.
  - Fortalecimento da inovação e colaboração interinstitucional.
- **Próximos Desafios**
  - Expandir as capacidades analíticas e garantir a segurança dos dados.
  - Aprofundar a integração de novas fontes para responder às necessidades futuras.





INSTITUTO NACIONAL DE ESTATÍSTICA  
STATISTICS PORTUGAL

# Infraestrutura Nacional de Dados

## Desenvolvimentos recentes