


# Padrões de inovação na indústria de transformação brasileira: uma análise exploratória a partir de dados da PINTEC semestral 2021

Marco Antonio VARGAS<sup>1</sup>

mvargas@id.uff.br  <https://orcid.org/0000-0001-7954-8594>


Rafael Ribeiro MACHARETE<sup>1</sup>

rafael\_macharete@id.uff.br  <https://orcid.org/0009-0000-1551-7369>


Jorge Nogueira de Paiva BRITTO<sup>1</sup>

jorgebritto@id.uff.br  <https://orcid.org/0000-0003-0730-4958>

João Carlos FERRAZ<sup>2</sup>

jcferraz@ie.ufrrj.br  <https://orcid.org/0000-0001-5980-2591>

Marina Honório de Souza SZAPIRO<sup>2</sup>

marina@ie.ufrrj.br  <https://orcid.org/0000-0003-4549-1192>

## Resumo

Este artigo analisa os padrões de esforço e desempenho inovativo na indústria de transformação brasileira a partir dos dados da Pesquisa de Inovação Tecnológica (PINTeC) semestral de 2021. A análise, de caráter exploratório, propõe uma metodologia para clusterização de setores mediante o uso de técnicas de aprendizado de máquina não supervisionado. O exercício resultou na identificação de quatro agrupamentos de setores da indústria de transformação que apresentam padrões diferenciados de esforço e desempenho inovativo, particularmente no tocante à sua aderência com as taxonomias de intensidade tecnológica e padrões setoriais já consagradas na literatura internacional.

## Palavras-chave

Padrões de Inovação, Indústria de transformação, Pesquisa de inovação, Aprendizado de máquina não supervisionado.

<sup>1</sup> Universidade Federal Fluminense, Faculdade de Economia, Niterói, RJ Brasil.

<sup>2</sup> Universidade Federal do Rio de Janeiro, Instituto de Economia, Rio de Janeiro, RJ, Brasil.

Recebido: 23/08/2023.

Revisado: 07/10/2024.

Aceito: 23/12/2024.

DOI: <https://doi.org/10.1590/1980-53575525mrjjm>



This Article is Distributed Under the Terms of the Creative Commons Attribution 4.0 International License.

# Innovation patterns in the brazilian manufacturing industry: an exploratory analysis based on data from the 2021 semiannual PINTEC survey

## Abstract

This article analyzes the patterns of effort and innovative performance in the Brazilian manufacturing industry. The analysis, based on data from the 2021 Semiannual Pesquisa de Inovação Tecnológica (PINTEC), proposes a methodology for clustering sectors using unsupervised machine learning techniques. The results allow to identify four groupings of sectors in the manufacturing industry that present different patterns of effort and innovative performance, particularly regarding their adherence to taxonomies of technological intensity and sectorial innovation patterns already enshrined in the international literature.

## Keywords

Innovation patterns, Manufacturing industry, Innovation surveys, Unsupervised machine learning.

## JEL Classification

C81; O31; O32.

## 1. Introdução

O principal objetivo do artigo reside na identificação de padrões diferenciados de desempenho e esforço inovativo em grupos de setores da indústria de transformação brasileira, a partir de uma proposta alternativa de agregação setorial. Com base num conjunto de 21 variáveis selecionadas da PINTEC Semestral 2021, o artigo propõe uma metodologia para clusterização de setores mediante o uso de técnicas de aprendizado de máquina não supervisionado.

A agregação setorial a partir de critérios relacionados à dinâmica produtiva e tecnológica das empresas apresenta vantagens no sentido de facilitar o processamento e análise de um volume extenso de informações e simplificar a análise de estruturas produtivas complexas a partir da identificação de categorias representativas para fins de elaboração de políticas públicas (Cavalcante 2014).

Existem diferenças significativas na taxa e na organização das atividades inovativas entre os setores industriais. Em algumas indústrias a mudança técnica ocorre em um ritmo rápido, enquanto em outras ela é lenta e gradual; em algumas indústrias a inovação é realizada por um pequeno número de atores, enquanto em outras ela é distribuída por uma população mais ampla de empresas; em determinadas atividades a inovação está atrelada a elevados gastos em Pesquisa e Desenvolvimento (P&D) “in house” enquanto em outros prevalece o papel da difusão, adaptação e absorção de tecnologias geradas externamente ao setor (Malerba 2007). Isso sugere que apesar da alta variação das atividades de inovação no nível da empresa, cada setor mostra padrões específicos de comportamento, dando origem a modos de inovação setoriais. Dado que os padrões de mudança técnica, inovação e desempenho econômico são específicos do setor e muito diversos entre os setores, alguns autores (por exemplo, Robson *et al.* 1988; Edquist 1997; Malerba 2002; Malerba e Orsenigo 1997) destacam a pertinência de uma perspectiva setorial da abordagem da inovação a fim de desenvolver políticas de inovação mais bem orientadas. No âmbito de cada setor, a cumulatividade do conhecimento, sua apropriabilidade e condições de oportunidade, os quais constituem aspectos centrais do que a literatura denomina de “regimes tecnológicos” (Malerba e Orsenigo 1997), operam como parâmetros básicos que afetam o comportamento das firmas, com nítidas especificidades setoriais (Peneder 2007). Em contrapartida, a própria definição dos diversos “tipos” de inovação também tende a ser progressivamente refinada, para captar a diversidade dessas soluções no contexto de cada setor (Gault 2018; OCDE/EUROSTAT 2018).

Os padrões setoriais de inovação, no entanto, não são independentes da situação nacional ou supranacional. Hinloopen (2003), por exemplo, constata que o ambiente de inovação de um país afeta significativamente a eficiência com que os esforços de inovação são transformados em produtos novos ou aprimorados. Castellaci (2008) e Castellaci (2009) explica as diferenças entre países no desempenho da inovação de setores ordenados de acordo com a classificação setorial de Pavitt (1984) como sendo determinadas por diferenças nas políticas de inovação e situação econômica. Desse modo, o comportamento inovativo de uma indústria específica também varia entre os países. O desempenho da inovação em cada país depende da interação entre as características da economia nacional, o Sistema Nacional de Inovação, as políticas fiscais, as instituições do mercado de trabalho (dentre outras variáveis) com os Sistemas Setoriais de Inovação (SIS) e o perfil de especialização nacional.

A identificação de padrões setoriais de inovação permite, portanto, investigar a existência de diferenças nas características tecnológicas e econômicas de setores da indústria, ao mesmo tempo que permite identificar condutas comuns a empresas de um mesmo setor, ou mesmo similaridades entre os padrões e regimes tecnológicos entre setores, de modo a aprofundar a compreensão sobre as especificidades e peculiaridades de sistemas setoriais e tecnológicos de inovação (Pavitt 1984; Malerba 1992).

A análise apresentada neste artigo busca trazer uma contribuição metodológica de caráter exploratório, e parte da articulação de duas questões que têm orientado a construção recente de taxonomias de padrões setoriais para a indústria brasileira. A primeira está relacionada com a aderência dos padrões observados no Brasil aos existentes nos países desenvolvidos, captados a partir de diferentes taxonomias, a maior parte delas construídas a partir do trabalho original de Pavitt (1984). Essa questão pode sugerir indiretamente que existe no Brasil uma dinâmica setorial relativamente similar à daqueles países. A segunda refere-se à hipótese de que os países em desenvolvimento apresentam características específicas em sua estrutura produtiva e, portanto, os indicadores geralmente utilizados para a construção das taxonomias internacionalmente reconhecidas não são adequados para a compreensão dos padrões setoriais de inovação vigentes nesses países (Marques, Roselino e Mascarini 2019).

O artigo encontra-se organizado em quatro seções além desta introdução. A próxima seção traz um breve panorama sobre as duas principais classificações tecnológicas presentes na literatura, a taxonomia de Pavitt (1984) e a de intensidade tecnológica da OCDE (2011), e discute o alcance dessas taxonomias tecnológicas e setoriais tradicionais para identificação dos padrões de inovação na indústria de transformação no contexto dos países em desenvolvimento, a partir de diferentes contribuições que analisam a experiência brasileira. A terceira seção discute os aspectos metodológicos do estudo. Inicialmente, essa seção traz uma breve discussão sobre as mudanças conceituais e metodológicas presentes na PINTEC Semestral, tendo em vista a sua adequação à quarta edição do Manual de Oslo (2018). A seguir, a seção apresenta os procedimentos metodológicos que orientaram um exercício de agregação setorial desenvolvido a partir dos resultados da PINTEC Semestral 2021, e do uso de ferramentas matemáticas baseadas em aprendizado de máquinas. A quarta seção traz a análise dos resultados da PINTEC Semestral 2021 a partir dos agrupamentos de setores resultantes da aplicação da metodologia proposta. Adicionalmente, a seção analisa a

existência de convergências ou divergências entre o agrupamento de setores resultante da aplicação da metodologia de clusterização, a classificação de intensidade tecnológica da OCDE (2011) e a taxonomia de setores proposta por Pavitt (1984). A quinta seção sintetiza os principais resultados do estudo.

## 2. Limites e perspectivas das taxonomias tecnológicas e setoriais

Alguns objetivos gerais norteiam a construção de taxonomias setoriais de inovação. Em primeiro lugar, elas objetivam capturar as formas únicas em que a inovação ocorre em diferentes setores, reconhecendo que os fatores impulsionadores (papel da demanda, oportunidades tecnológicas, estruturas regulatórias e dinâmica da concorrência, dentre outros), os esforços e os resultados da inovação variam de acordo com o setor. Em segundo lugar, tais taxonomias devem permitir identificar diferenças entre setores a partir de dados robustos e pesquisas empíricas estruturadas. Em terceiro lugar, tais taxonomias devem permitir mapear fluxos de conhecimento setoriais e fontes de inovação, descrevendo como o conhecimento é gerado, compartilhado e aplicado dentro e entre setores, incluindo o papel das universidades, institutos de pesquisa, fornecedores e clientes no processo de inovação. Em quarto lugar, devem permitir acompanhar como os setores evoluem ao longo do tempo, particularmente em termos de trajetórias tecnológicas e taxa de adoção de inovação, articulando perspectivas globais e locais, evidenciando como a competição internacional, a difusão tecnológica e as condições locais e globais moldam os padrões de inovação setoriais. Por fim, destaca-se o caráter potencialmente normativo dessas taxonomias, fornecendo *insights* claros sobre como diferentes setores podem se beneficiar de políticas específicas, bem como informando decisões de investimento públicas e privadas, destacando oportunidades de crescimento e inovação em diferentes setores.

A classificação de setores por grau de intensidade tecnológica proposta pela Organização para a Cooperação e Desenvolvimento Econômico (OCDE 2011) e a classificação proposta por Pavitt (1984), que considera os padrões setoriais de mudança técnica, são taxonomias já consagradas na literatura e amplamente utilizadas tanto no meio acadêmico como por formuladores de políticas. Entretanto, para países em desenvolvimento, que geralmente contam com uma estrutura industrial que opera em patamares muito inferiores aos das fronteiras tecnológicas setoriais mundiais, observa-se a

necessidade de construção de classificações capazes de identificar as particularidades da sua dinâmica industrial e tecnológica (Silva e Suzigan 2014; Yonamini 2011; Campos e Urraca-Ruiz 2009; Furtado e Carvalho 2005).

A taxonomia de Pavitt tem origem no esforço de compreensão dos padrões setoriais de mudança tecnológica e busca explicar as semelhanças e diferenças entre setores em termos de fontes de informação, porte, atividades produtivas e setores de produção e uso das inovações adotadas por empresas inovadoras (Pavitt 1984). A taxonomia teve como base empírica um estudo sobre as características de cerca de 2.000 inovações significativas e de empresas inovadoras na Grã-Bretanha de 1945 a 1979 (Townsend *et al.* 1981). Originalmente, o autor propõe quatro categorias ou grupos de setores. Um primeiro grupo, de setores **dominados por fornecedores**, caracteriza-se pela menor intensidade dos gastos em P&D, inovações de caráter mais incremental voltadas à redução de custos, produtos mais padronizados e concorrência pautada mais por qualidade e preço, com foco em marcas, design e marketing. Um segundo grupo, congrega os chamados setores **intensivos em escala** que envolve o domínio de sistemas produtivos complexos e com foco em inovações incrementais de produto e de processo voltadas à melhoria do desempenho. Um terceiro grupo de setores, de **fornecedores especializados ou difusores de progresso técnico**, também constitui o que Pavitt (1984) denomina de setores de produção intensiva, mas se diferencia dos setores intensivos em escala por serem responsáveis por inovações que são inseridas nos demais setores como insumos de capital. As empresas neste grupo contam com conhecimentos especializados e de caráter tácito e abrange os produtores de máquinas e equipamentos e instrumentos (Marques *et al.* 2019). Por fim, o grupo de setores **baseados em ciência** é aquele onde as fontes de informação para a inovação são as atividades de P&D baseadas no papel das instituições ligadas à infraestrutura de ciência, tecnologia e inovação.

Apesar das críticas pelo seu caráter estático e das diversas tentativas de revisão, a taxonomia proposta por Pavitt (1984) constitui uma das mais importantes contribuições no estudo dos processos de mudança tecnológica e segue sendo amplamente utilizada (Archibugi 2001).

A classificação de setores por intensidade tecnológica foi registrada originalmente por Hatzichronoglou (1997), a partir de critérios que procuram levar em conta tanto o nível de tecnologia específica do setor (medido pela razão entre os gastos em P&D e o valor adicionado) quanto a tecnologia

incorporada nas compras de bens intermediários e de capital (Cavalcante 2014). A classificação da OCDE corresponde a uma agregação de setores econômicos a partir de um método mais objetivo de agregação na medida em que focaliza a relação entre os gastos em P&D e o valor adicionado, ou na relação entre os gastos em P&D e o faturamento. A classificação já sofreu diversas revisões (OECD 2007 e European Commission 2008), sendo a mais recente de 2011 (OCDE 2011), e contempla quatro categorias de setores: i) setores de alta intensidade em P&D; ii) setores de média-alta intensidade em P&D; iii) setores de média-baixa intensidade em P&D e; (iv) setores de baixa intensidade em P&D. Cabe ressaltar, conforme destacado por Cavalcante (2014), que a taxonomia da OCDE não pressupõe a existência de homogeneidade nos padrões setoriais de mudança técnica nos setores que integram cada um dos grupos por intensidade tecnológica. Dessa forma, não existe uma pretensão de que os grupos respondam de forma homogênea aos estímulos de políticas e ações específicas voltadas a ampliar o dinamismo inovativo das firmas.

As taxonomias setoriais de inovação, apesar da sua funcionalidade para categorizar padrões, apresentam certas lacunas reconhecidas pela literatura. Muitas taxonomias de inovação são fortemente focadas em setores da manufatura, frequentemente negligenciando serviços e outros setores como agricultura, serviços públicos e indústrias criativas, apesar dos esforços para “estender” essas taxonomias para além da manufatura (Castellacci 2008). Em segundo lugar, as taxonomias tradicionais frequentemente falham em levar em conta a crescente convergência de setores, em particular em função do impacto transversal e intersetorial das “tecnologias de propósito geral” (como as tecnologias de base digital). Em terceiro lugar, cabe mencionar a natureza estática das taxonomias, que frequentemente falham em capturar a natureza dinâmica da inovação, podendo levar a classificações desatualizadas ou imprecisas, dificultando a formulação de políticas eficazes e a gestão da inovação. Em quarto lugar, as taxonomias tradicionais podem não levar em conta suficientemente as inovações sociais ou voltadas para a sustentabilidade, inclusive devido a uma sub-representação dos setores informais e emergentes. Por fim, estas taxonomias podem não refletir adequadamente a natureza globalizada da inovação, onde as indústrias operam através de fronteiras e cadeias de valores globais, podendo resultar em uma compreensão incompleta de como as redes globais e a colaboração internacional influenciam os processos de inovação.

Além dessas limitações de natureza geral, o uso de tais taxonomias, construídas originalmente a partir da realidade industrial e tecnológica observada nos países desenvolvidos há décadas, apresenta limitações para análise dos padrões setoriais de inovação observados nos países em desenvolvimento. Dentre os principais argumentos que buscam fundamentar a análise sobre as especificidades inerentes aos padrões setoriais de inovação nos países em desenvolvimento, destacam-se a elevada heterogeneidade estrutural da base produtiva industrial, a existência de padrões de especialização que refletem formas subordinadas de inserção na divisão internacional do trabalho ou a existência de assimetrias tecnológicas em relação à fronteira tecnológica onde operam os setores industriais nos países desenvolvidos (Cimoli e Porcile 2011; Marques, Roselino e Mascarini 2019).

A heterogeneidade estrutural de países em desenvolvimento limita severamente sua capacidade inovadora ao criar grandes lacunas em termos de produtividade, habilidades e suporte institucional em diferentes setores. A concentração de inovação em alguns setores avançados, a fraca integração entre setores formais e informais e a falta de estruturas institucionais abrangentes dificultam a difusão da inovação por toda a economia. No plano tecnológico, as empresas em países em desenvolvimento frequentemente operam longe da fronteira tecnológica e podem focar em inovação adaptativa, imitação ou melhorias incrementais em vez de serem pioneiras em novas tecnologias. A ênfase em padrões de inovação baseados em P&D avançado pode negligenciar o papel da difusão, adaptação e absorção de tecnologia, que são mais críticas nessas regiões.

As taxonomias setoriais de inovação geralmente pressupõem mercados sofisticados e de alta renda que exigem inovações de ponta e têm a capacidade de financiar e apoiar tais inovações. Em contraste, muitos países em desenvolvimento enfrentam limitações do lado da demanda, com menor poder de compra, acesso limitado aos mercados e foco em inovações frugais e de baixo custo. Essas realidades de mercado moldam os padrões de inovação de forma diferente, muitas vezes priorizando a redução de custos e a acessibilidade em vez de avanços tecnológicos. É importante considerar também que os países em desenvolvimento estão frequentemente em processo de industrialização, o que significa que, comparativamente aos países desenvolvidos, podem seguir trajetórias de inovação muito diferentes, inclusive incorporando estratégias de salto (*leapfrogging*), não adequadamente refletidas naquelas taxonomias.



Arocena e Sutz (2000), identificam especificidades das indústrias de países em desenvolvimento, no que se refere a geração de conhecimento, fontes de informação, aprendizado e inovação, diferenciando-as dos padrões encontrados em países desenvolvidos, o que revelaria uma inadequação da utilização daqueles padrões como ferramenta para a análise de estruturas industriais singulares e heterogêneas. Além disso, dado o padrão de inserção dos países em desenvolvimento nos fluxos internacionais de comércio – ou nas cadeias globais de valor – a competitividade da indústria desses países não necessariamente decorre da capacidade de um determinado setor se posicionar na fronteira tecnológica, e sim, dentre outros aspectos, da produtividade e do custo de fatores de produção críticos, o que pode influenciar a construção de taxonomias industriais adequadas àquelas realidades (Hermida e Xavier 2012). Além disso, conforme ressaltado por Furtado e Carvalho (2005), nos países em desenvolvimento, as economias são mais fechadas, e, geralmente, os setores de maior intensidade tecnológica estão voltados para o mercado interno, limitando o acesso ao fluxo de conhecimento internacional. Por fim, a organização estrutural e institucional particular dos mercados, em conjunto com a dimensão geográfica e sócio populacional, também podem interferir nas características dos setores industriais, dificultando a transposição de taxonomias estilizadas para aquelas realidades.

É no contexto desse debate que se inserem estudos que procuram analisar a aderência do padrão setorial da inovação da indústria brasileira aos padrões dos países desenvolvidos. Duas linhas de abordagem se destacam nessa direção.

Por um lado, alguns estudos procuram analisar em que medida recortes tradicionais de diferenciação setorial com base em indicadores de “intensidade tecnológica” poderiam ser aplicados à realidade brasileira. Furtado e Carvalho (2005), por exemplo, buscaram comparar dados de dispêndio e de recursos humanos em P&D do Brasil com os países da OCDE. O estudo destaca diferenças estruturais significativas nos padrões setoriais de esforço tecnológico, apontando para esforços tecnológicos mais limitados no Brasil do que nos países desenvolvidos, principalmente em setores de “alta intensidade tecnológica”, evidenciando um quadro geral de maior homogeneidade e de baixa intensidade do esforço tecnológico no Brasil quando comparada a países desenvolvidos. Moura *et al.* (2015) procuram classificar os setores da indústria nacional por níveis de intensidade tecnológica, considerando não apenas a intensidade em pesquisa e desenvolvimento, mas também um indicador da intensidade em práticas de inovação, iden-

tificando uma certa defasagem com relação à magnitude de intensidade de investimento em tecnologia, com uma grande concentração dos setores industriais em níveis baixos de tecnologia, com pequeno esforço em apropriação de tecnologia externa e criação de tecnologia própria. Kannebley Junior *et al.* (2005) utilizam procedimentos estatísticos não-paramétricos para identificar os principais fatores distintivos entre empresas inovadoras e não-inovadoras, destacando a orientação exportadora, o tamanho da empresa e a origem estrangeira do capital; em seguida, ampliam a análise para o plano interindustrial, através da caracterização das condições de “oportunidade tecnológica” para 21 setores industriais.

Por outro lado, outra linha refere-se, justamente, à construção de taxonomias com base em indicadores de esforços, práticas, condicionantes e desempenho inovativo e à verificação da maior ou menor aderência dessa taxonomia em relação a um “quadro analítico” de referência, em geral a classificação de setores segundo “padrões tecnológicos” elaborada por Pavitt (1984). Algumas dessas análises se mostram mais “arraigadas” a este quadro analítico de referência.

O estudo de Campos e Ruiz (2009), a partir da base de dados da PINTEC 2000, faz uso da análise de clusters, identificando que o perfil inovativo dos setores industriais brasileiros, apesar de apresentar alguma coerência com as proposições do modelo pavittiano, apresenta particularidades em setores específicos. O trabalho de Castro (2010) analisa o padrão de inovação tecnológica nos setores industriais brasileiros a partir da PINTEC 2005, tomando como referência o modelo de taxonomia pavittiana. Utilizando a combinação dos métodos hierárquicos e não-hierárquicos, a análise define quatro clusters, a partir dos quais são classificados cerca de vinte e três setores da indústria brasileira, correspondendo a setenta e sete por cento do total dos trinta setores estudados, porém sem que nenhum setor tenha apresentado um padrão claro de inovação tecnológica pavittiano do tipo fornecedores especializados.

Hermida e Xavier (2012) discutem o desempenho e a competitividade das exportações brasileiras de acordo com a taxonomia de Pavitt, enfatizando o comportamento de setores com maior grau tecnológico, porém adaptando o recorte setorial original a um recorte vinculado à intensidade de recursos, conforme proposto por Laplane *et al.* (2001). Silva e Suzigan (2014) também chegam à conclusão de que a indústria de transformação brasileira revela uma adequação com a taxonomia de Pavitt, porém com

qualificações e exceções importantes, refletindo tanto a heterogeneidade estrutural setorial da indústria brasileira, como peculiaridades da formação econômica brasileira que resultaram em comportamentos distintos dos padrões internacionais, destacando-se setores que conciliam baixa densidade tecnológica e elevada participação das exportações na geração da receita, os quais apresentam baixa aderência às suas respectivas categorias pavittianas.

Marques, Roselino e Mascarini (2019) utilizam a PINTEC combinada com dados da PIA-Empresa e RAIS, também utilizando uma análise de cluster, concluindo no sentido de uma relativa aderência aos padrões internacionais com exceção da atividade de refino de petróleo, que geralmente é classificada como média-baixa tecnologia, mas no Brasil foi classificada como alta tecnologia. Por outro lado, atividades que geralmente são classificadas como de alta intensidade tecnológica, em geral estão muito aquém disso no Brasil, o que poderia ser explicado pela presença de multinacionais estrangeiras que concentram as atividades de P&D nos países centrais. A análise de Lucena (2022) também utiliza dados da PINTEC complementados com dados da RAIS e PIA, combinando análise fatorial para encontrar fatores relevantes e uma análise de cluster para criar grupos setoriais com as mesmas características de mudança tecnológica. O artigo sugere que a indústria brasileira apresenta alguma convergência com a taxonomia de Pavitt, mas apresenta algumas especificidades que também foram identificadas em estudos anteriores. A indústria brasileira ainda concentra seus setores em grupos com baixo dinamismo inovativo, que gastam pouco com inovação e, quando gastam, o fazem apenas para pequenos projetos de baixo impacto inovativo. O artigo também sugere que o porte da empresa não é uma boa variável para investigar o padrão setorial da indústria brasileira.

Assim, apesar da relevância da taxonomia de Pavitt para o entendimento do comportamento e da trajetória de inovação dos setores industriais e como ferramenta para a tomada de decisão por gestores de políticas em economias desenvolvidas, as evidências sobre as características específicas da estrutura industrial de países em desenvolvimento indicam não ser possível uma total aderência a tal taxonomia, tanto em função da heterogeneidade inter e intrasetorial como do progresso técnico lento e desigual.

Cabe, portanto, questionar as taxonomias industriais essencialmente rígidas e fortemente embasadas na realidade de países desenvolvidos como ferramenta analítica para a compreensão da realidade de países em desenvolvimento (Marques, Roselino e Mascarin 2019). Em particular, é importante considerar que alguns setores industriais brasileiros apresentam particularidades que devem ser bem explicitadas para embasar a elaboração de políticas industriais sistêmicas e eficazes. Neste contexto, observa-se a necessidade de encontrar taxonomias nacionais menos “rígidas”, que incorporem as especificidades, a dinâmica industrial e a trajetória tecnológica dos setores para cada país.

### 3. Aspectos metodológicos

#### 3.1. PINTEC Semestral 2021: desafios conceituais e metodológicos

A PINTEC Semestral 2021 representa um esforço inédito e importante na construção de uma pesquisa conjuntural de inovação no Brasil. A pesquisa permite a produção e divulgação de estatísticas relevantes sobre o tema da inovação nas empresas, de forma ágil e com periodicidade menor do que as pesquisas de inovação estruturais, como é caso da PINTEC trienal (IBGE 2022). Por se tratar de uma pesquisa nova e de caráter experimental, a PINTEC Semestral apresenta desafios conceituais, metodológicos e analíticos importantes.

As pesquisas de inovação conjunturais se destacam pela sua capacidade de prover informações sobre mobilização de recursos, esforço e desempenho inovativo de empresas de forma recorrente e ágil. O caráter recorrente destas pesquisas (que podem ser trimestrais, semestrais ou anuais) constitui a sua principal vantagem na medida em que permite a disponibilização de indicadores com maior frequência, o acompanhamento das condutas inovativas das empresas, bem como dos fatores que afetam e condicionam tais condutas. Tal conhecimento é importante não somente para qualificação e calibragem das pesquisas de inovação estruturais, como também para melhoria dos instrumentos de avaliação das políticas públicas de inovação e para a definição de estratégias empresariais (Ferraz *et al.* 2022).

A recorrência da pesquisa e a possibilidade de comparação com outras pesquisas e comparação intertemporal de indicadores semelhantes, garantida pela estabilidade da amostra, é o que valoriza as pesquisas de inovação de curto prazo. O caráter recorrente (trimestral, semestral ou anual) das pesquisas tende a orientar, em grande parte, a formulação de questões em termos do número (limitado) e da especificação (simples) das questões. Essa recorrência influencia a definição do painel de empresas e o perfil dos respondentes. Também é importante considerar a natureza das demandas dos potenciais usuários das Pesquisas de Inovação (academia, empresas, formuladores de políticas etc.). Esses indicadores podem incluir informações quantitativas e qualitativas, gerando métricas (individuais ou compostas) que fornecem uma representação sintética do fenômeno da inovação.

Entretanto, o caráter recorrente das pesquisas conjunturais tende a limitar o alcance do universo de empresas pesquisadas em termos do seu número e perfil, grau de detalhamento e número de questões a serem incorporadas no levantamento de informações. Os indicadores devem ser simples e de fácil compreensão, possíveis de serem informados e avaliados sistematicamente (Ferraz *et al.* 2022; Ferraz e Britto 2020).

A PINTEC Semestral é uma pesquisa empírica semestral, que envolve a coleta de dados de empresas industriais sediadas no Brasil por meio da aplicação de questionário eletrônico por entrevistadores treinados, sob supervisão de especialistas, realizado pelo Instituto Brasileiro de Geografia e Estatística (IBGE), órgão oficial de produção de informações estatísticas do governo brasileiro. O questionário eletrônico foi desenvolvido especificamente para a pesquisa, sendo atualizado a cada ciclo de sua edição. O instrumento de coleta está integrado a um sistema de banco de dados baseado em Tecnologias de Informação e Comunicação (TIC) que permitem a captação, registro, processamento e disponibilização dos dados coletados por meio de relatórios e tabulações, permitindo o gerenciamento da coleta de dados e seu fluxo de processamento. A pesquisa sobre indicadores básicos a que se refere este artigo resulta de um questionário aplicado no segundo semestre de cada ano, cujo período de referência é o ano anterior, tendo como tema central a inovação e o esforço de P&D.

A fim de alcançar o objetivo de ser uma pesquisa ágil que fornece informações atualizadas sobre o estágio da inovação nas empresas industriais brasileiras e subsidiar a formulação e implementação de políticas públicas de apoio à inovação empresarial, a PINTEC Semestral apresenta um escopo menor tanto em termos de indicadores como em relação ao período de abrangência e o universo de empresas contempladas na pesquisa. Nesse sentido, a pesquisa considera um período de referência anual para os dados divulgados, abarca um universo de empresas com 100 ou mais pessoas ocupadas e possui menor nível de desagregação setorial. O seu escopo menos abrangente, em comparação à PINTEC Trienal, é parcialmente compensado pelo caráter mais focado da amostra, que ao se limitar a um universo de empresas com mais de 100 empregados, minimiza as diferenças intersetoriais. Tais questões revelam a natureza diferenciada e inédita da PINTEC Semestral e, ao mesmo tempo, impõem cautela no estabelecimento de análises comparativas com as edições anteriores da PINTEC trienal.

Igualmente importante é o fato de que a PINTEC Semestral incorpora mudanças conceituais da quarta edição do Manual de Oslo (OCDE 2018). Uma das principais diferenças refere-se à definição de inovação de processo que passa a ser designada de “inovações de processos de negócios”, contemplando também as inovações organizacionais e de marketing, com um caráter mais amplo do que na definição das edições anteriores da PINTEC estrutural. Neste aspecto, ao incorporar o conceito de inovações em processos de negócios a PINTEC Semestral traz não somente uma ampliação do conceito de inovação, como aproxima a importância relativa entre inovações de produtos (incluindo bens e serviços), processos, marketing e organizacional (Ferraz *et al.* 2022). Adicionalmente, a pesquisa incorpora mais informações com vistas à qualificação das expectativas das empresas e do ambiente de inovação onde estas operam.

A pesquisa traz pelo menos quatro aspectos diferenciados e importantes, no tocante à avaliação de esforços e desempenho inovativo das empresas:

- a) Uma visão mais abrangente dos tipos de inovação de produto e processos de negócios;
- b) Uma preocupação com aspectos do “ambiente” (interno e externo) no qual os esforços inovativos são realizados, refletidos tanto na identificação de obstáculos como no seu enfrentamento através de práticas cooperativas;

- c) A preocupação em incorporar a evolução de expectativas acerca da intensidade e continuidade dos esforços inovativos; e
- d) A preocupação em incorporar elementos de caracterização mais detalhada das empresas inovadoras, como porte, setor e origem de capital.

### 3.2. *Análise de clusters*

A análise dos padrões de esforço e desempenho inovativo a partir do agrupamento de setores envolveu a elaboração de 22 indicadores relativos aos 24 setores de atividades econômicas da indústria de transformação contemplados nos resultados da PINTEC Semestral 2021. Os indicadores escolhidos contemplam variáveis que captam os resultados das inovações introduzidas e o esforço inovativos empreendido pelas firmas. Em termos dos resultados foram contempladas a taxa geral de inovação e as taxas de inovação de produto, seu grau de novidade, e as taxas de inovação em processos de negócios. Em termos das variáveis de esforço inovativo foi contemplado o percentual de empresas que realizaram dispêndios em atividades de P&D interno e as relações de cooperação com outras organizações. Em razão de limitações metodológicas e estatísticas da PINTEC Semestral 2021, que impossibilitam o uso de microdados coletados ao nível da firma, não foi possível a incorporação na análise de informações de outras bases, como a PIA-IBGE, relativas à estrutura e desempenho industrial, ficando a análise restrita aos dados de esforço e desempenho inovativo disponíveis no plano tabular padrão da PINTEC Semestral. O Quadro 1 mostra os 22 indicadores selecionados. O detalhamento destes indicadores para os 24 setores da indústria de transformação encontra-se no anexo 1 do artigo.

Quadro 1 - Indicadores da PINTEC

Variável	Descrição
$X_1 = \text{TX-INOV}$	Taxa Geral de inovação
$X_2 = \text{TX inov PROD}$	Inovação em produto.
$X_3 = \text{TX inov PROD-EMP}$	Inovação em produto para empresa.
$X_4 = \text{TX inov PROD-NAC}$	Inovação em produto para mercado nacional.
$X_5 = \text{TX inov PROD-MUN}$	Inovação em produto para mercado mundial.
$X_6 = \text{TX inov PROC}$	Inovação em processos de negócios.
$X_7 = \text{InovPROC-PROD}$	Inovação na produção de bens ou fornecer serviços.
$X_8 = \text{InovPROC-LOG}$	Inovação na logística, entrega ou distribuição.
$X_9 = \text{InovPROC-TIC}$	Inovação no processamento de informação ou comunicação
$X_{10} = \text{InovPROC-CONT}$	Inovação em contabilidade ou outras operações administrativas
$X_{11} = \text{InovPROC-GEST}$	Inovação nas Práticas de Gestão para organizar procedimentos organizativos ou relações externas
$X_{12} = \text{InovPROC-ORG}$	Inovação na organização do trabalho, de tomada de decisão ou de gestão de recursos humanos
$X_{13} = \text{InovPROC-MKT}$	Inovação em marketing para promoção, embalagem, preços, colocação de produtos ou serviços pós-venda
$X_{14} = \text{Inov-COOP}$	Inovação em cooperação
$X_{15} = \text{Inov-COOP-CLIEN}$	Inovação em cooperação com clientes.
$X_{16} = \text{Inov-COOP-FORN}$	Inovação em cooperação com fornecedores.
$X_{17} = \text{Inov-COOP-CONC}$	Inovação em cooperação com concorrentes.
$X_{18} = \text{Inov-COOP-CT\&I}$	Inovação em cooperação infraestrutura de ciência, tecnologia e inovação.
$X_{19} = \text{Inov-COOP-STRUP}$	Inovação em cooperação com Start-Ups.
$X_{20} = \text{Inov-COOP-CONS}$	Inovação em cooperação com consultores e empresas de consultoria.
$X_{21} = \text{Inov-COOP-OEMP}$	Inovação em cooperação com outras empresas do grupo.
$X_{22} = \text{Inov-DISP}$	Inovação com dispêndio em Pesquisa e desenvolvimento

Fonte: Elaboração própria a partir da PINTEC Semestral 2021.

A realização dos agrupamentos, foi feita através da técnica de aprendizado de máquinas não supervisionado. O aprendizado de máquinas é uma área da computação que utiliza métodos estatísticos, inteligência artificial e ciência da computação para identificar padrões em dados (Tan e Kumar 2014 e Rezende 2005). No aprendizado de máquinas não supervisionado, não é necessário indicar respostas para o algoritmo, permitindo que o agrupamento seja realizado somente com base nas características dos dados (Alpaydin 2010). O *clustering* é um método de particionamento de dados em grupos com base na similaridade dos mesmos (Jain 2010).



Na análise, foi utilizado o algoritmo K-means (Hartigan e Wong 1979), que trata os dados em um espaço de  $n$  dimensões correspondentes aos indicadores. O algoritmo determina, de forma recorrente, o centro dos grupos, chamado de centróide, e a cada iteração, os pontos são realocados para a centróide mais próxima, calculada pela média dos pontos daquela região. O algoritmo é finalizado quando não há mais alterações nos grupos, e é dito que o algoritmo convergiu (Guido 2016).

O Quadro 2 apresenta a matriz de correlação para os pares de variáveis analisadas. Observando a matriz, pode-se perceber que há fortes correlações positivas entre algumas variáveis. Por exemplo, há correlação positiva forte entre as variáveis “taxa Geral de inovação” (TX inov) e “inovação com dispêndio em Pesquisa e desenvolvimento” (Inov-DISP). Em geral, é possível observar que a maioria das correlações entre as variáveis de inovação é positiva, indicando que há uma associação entre essas variáveis. Já as correlações negativas mais fortes foram encontradas entre “inovação em produto para mercado mundial” (TX inov PROD-MUN) e “inovação em cooperação com clientes” (Inov-COOP-CLIEN). As variáveis “inovação na logística, entrega ou distribuição” (InovPROC-LOG) e “Inovação em contabilidade ou outras operações administrativas” (InovPROC-CONT) apresentaram correlações fracas com as demais variáveis de inovação, indicando que essas variáveis podem não estar tão associadas com as outras. Além disso, é possível observar que algumas variáveis possuem correlações com outras variáveis de outras classes, como as variáveis de cooperação “inovação em cooperação com clientes” (Inov-COOP-CLIEN) e “inovação em cooperação com fornecedores” (Inov-COOP-FORN), que apresentaram correlações moderadas a forte com a taxa de inovação em produto (TX inov PROD). Isso sugere que a inovação em produto pode estar relacionada à cooperação com clientes e fornecedores.



Para determinar o número ideal de grupos ou clusters a serem analisados, foi aplicado o método do cotovelo e avaliação da soma dos quadrados dentro dos grupos. Os estudos anteriores mencionados (Campos e Urraca 2009; Castro 2010; Silva e Suzigan 2014; Marques, Roselino e Marcarini 2019; Lucena, 2022) utilizam métodos de agrupamento estatísticos tradicionais da análise multivariada, como os métodos hierárquicos e não-hierárquicos, para definir o número de clusters e realizar a detecção de grupos. Embora robustos, esses métodos foram desenvolvidos em um período em que a disponibilidade e o tratamento de dados eram mais limitados do que os atuais. Com o avanço do Big Data e da capacidade computacional, essas técnicas foram aprimoradas para lidar com a crescente complexidade e volume de dados. A metodologia de aprendizado de máquina não supervisionado utilizada neste trabalho representa uma evolução nessa direção, introduzindo ferramentas da ciência da computação no campo da Análise Multivariada.

Nos trabalhos mencionados, a definição do número de clusters foi feita por meio de métodos hierárquicos, que geram dendrogramas para visualizar a formação dos clusters em diferentes níveis. O método do cotovelo, utilizado neste trabalho, é uma alternativa mais objetiva, avaliando a soma das distâncias dentro dos clusters para diferentes números de agrupamentos. Ele identifica o ponto (cotovelo) onde o ganho de eficiência na separação dos dados diminui significativamente, gerando uma curva que permite escolher o número ideal de clusters. O método do cotovelo é mais apropriado, pois otimiza a análise e evita a necessidade de inspeção visual, como nos métodos hierárquicos.

Do mesmo modo que nos trabalhos citados, a avaliação dos clusters utilizou medidas estatísticas de similaridade interclusters e dissimilaridade intraclusters. Também aplicamos dois modelos hierárquicos para avaliação dos clusters e realizamos uma análise algébrica, que calculou as distâncias entre os vetores dos setores dentro e entre os clusters. Em termos da fundamentação metodológica, o trabalho baseou-se em Dangeti (2017) para definição dos clusters, bem como em Fávero (2022) e Morettin (2023) para a detecção e avaliação dos clusters.

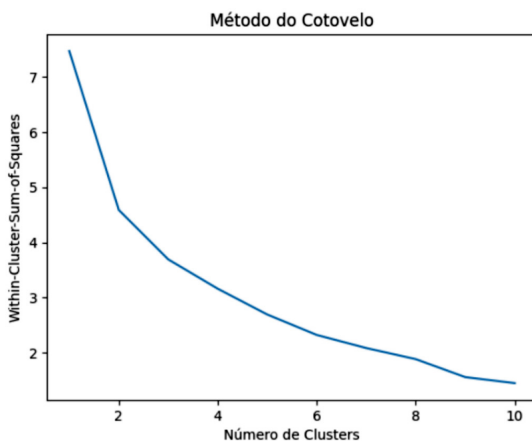


Figura 1 - Método do cotovelo aplicado nos quatro indicadores

Fonte: Elaboração própria.

O método do cotovelo gera uma função para diferentes valores de  $k$  e nos permite avaliar a soma dos quadrados dentro dos grupos para encontrar o melhor  $k$  (Dangeti 2017). Após a aplicação do método nos dados, verificou-se que o número ideal de clusters pode variar de dois a oito. No entanto, observa-se um baixo ganho na diferenciação de novos grupos a partir desses valores, como pode ser visto na Figura 1.

Com base nos resultados obtidos pelo método do cotovelo, foi implementado o algoritmo  $k$ -means para  $k$  igual a quatro e foram encontrados, inicialmente, os seguintes clusters, conforme apresentados no Quadro 3.

### Quadro 3 - Clusterização dos setores industriais

<b>Cluster 0</b>
Equipamentos de informática, produtos eletrônicos e ópticos
Produtos Farmoquímicos e farmacêuticos
<b>Cluster 1</b>
Coque, derivados do petróleo e biocombustíveis
Fumo
Manutenção, reparação e instalação máq equip
Metalurgia
Produtos de Madeira
<b>Cluster 2</b>
Couros e calçados
Máquinas e equipamentos
Máquinas, aparelhos e materiais elétricos
Produtos diversos
Produtos químicos
Veículos automotores, reboques e carrocerias
<b>Cluster 3</b>
Bebidas
Artigos de borracha e plástico
Celulose, papel e produtos de papel
Impressão e reprodução de gravações
Móveis
Outros equipamentos de transporte
Produtos alimentícios
Produtos de metal
Produtos de minerais não-metálicos
Produtos têxteis
Vestuário e acessórios

Fonte: Elaboração própria.

## 4. Padrões setoriais de inovação a partir do agrupamento de setores

### 4.1. Caracterização geral

A análise dos resultados da PINTEC Semestral a partir da metodologia de clusterização proposta na segunda seção permitiu a identificação de quatro agrupamentos de setores que apresentam padrões diferenciados em termos do conjunto de 21 variáveis selecionadas de esforço e desempenho inovativo.

O primeiro agrupamento identificado (**Cluster 0 – Setores intensivos em conhecimento**), conta com setores intensivos em conhecimento e é composto pelos setores de Fabricação de produtos farmoquímicos e farmacêuticos (Divisão CNAE 21) e Fabricação de equipamentos de informática, produtos eletrônicos e ópticos (Divisão CNAE 26). Este grupamento de setores apresenta a maior taxa geral de inovação, além de contar com o maior percentual de empresas com dispêndios em atividades de P&D e as maiores taxas de cooperação para inovação com diferentes tipos de atores, variáveis que diferenciam consideravelmente esse grupo dos demais.

O segundo agrupamento de setores (**Cluster 1 – Setores intensivos em escala**), congrega os setores de Fabricação de produtos do fumo (Divisão CNAE 12), Fabricação de produtos de madeira (Divisão CNAE 16), Fabricação de coque, produtos derivados do petróleo e de biocombustíveis (Divisão CNAE 19), Metalurgia (Divisão CNAE 24) e Manutenção, Reparação e Instalação de Máquinas e Equipamentos (Divisão CNAE 33). Este grupo é composto por setores de baixa e média-baixa intensidade tecnológica que apresentam as menores taxas de inovação e de esforço inovativo entre os setores analisados, particularmente no tocante à participação de empresas com dispêndio em atividades de P&D.

O terceiro agrupamento (**Cluster 2 – Setores baseados em processos industriais complexos**), congrega setores com processos industriais complexos e é composto pelos setores de Fabricação de produtos químicos (Divisão CNAE 20), Fabricação de máquinas, aparelhos e materiais elétricos (Divisão CNAE 27), Fabricação de máquinas e equipamentos (Divisão CNAE 28), Fabricação de veículos automotores, reboques e carrocerias (Divisão CNAE 29), Fabricação de produtos diversos (Divisão CNAE 32) e Preparação de couros e fabricação de artefatos de couro, artigos para viagem e calçados (Divisão CNAE 15). Este grupo é composto majoritariamente por setores de média-alta intensidade tecnológica, porém ele contempla a participação de dois setores com baixa intensidade tecnológica (calçados e produtos diversos) que apresentam taxas de esforço e desempenho inovativo compatíveis com o padrão dos demais setores do grupo. Cabe ressaltar que o Cluster 2 apresenta taxas de inovação de produto e processos de negócios muito próximas daquelas observadas no Cluster 0, mas se diferencia pelas menores taxas de cooperação para inovação e pela menor participação relativa de empresas com dispêndio em atividades de P&D.

O quarto e último agrupamento (**Cluster 3 – Setores dependentes de fornecedores**), é o que agrega o maior número de setores (11) e é composto pelas atividades de Impressão e reprodução de gravações (Divisão CNAE 18), Fabricação de produtos de metal (Divisão CNAE 25), Fabricação de produtos alimentícios (Divisão CNAE 10), Fabricação de bebidas (Divisão CNAE 11), Fabricação de produtos têxteis (Divisão CNAE 13), Confecção de artigos do vestuário e acessórios (Divisão CNAE 14), Fabricação de celulose, papel e produtos de papel (Divisão CNAE 17), Fabricação de artigos de borracha e plástico (Divisão CNAE 22), Fabricação de produtos de minerais não-metálicos (Divisão CNAE 23), Fabricação de outros equipamentos de transporte (Divisão CNAE 30) e Fabricação de móveis

(Divisão CNAE 31). O grupo é composto majoritariamente por setores de baixa e média-baixa intensidade tecnológica, ainda que conte com a participação de pelo menos um setor de média-alta intensidade tecnológica (Outros equipamentos de transporte). Este grupo de setores também é o que apresenta o menor desvio padrão em relação à média da indústria de transformação em todos os indicadores de esforço e desempenho inovativo. Além disso, ele se diferencia do Cluster 1, que congrega setores no mesmo grupo de intensidade tecnológica, por apresentar, em média, um melhor desempenho em todos os indicadores de esforço e desempenho inovativo.

O Quadro 4 apresenta uma síntese da caracterização dos quatro agrupamentos de setores identificados a partir da metodologia apresentada na terceira seção.

**Quadro 4 - Caracterização dos clusters de setores**

Descrição	Cluster 0 (2 setores)	Cluster 1 (5 setores)	Cluster 2 (6 setores)	Cluster 3 (11 setores)
	<b>Alta atividade inovativa - intensivos em conhecimento</b>	<b>Baixa atividade inovativa – intensivos em escala</b>	<b>Alta atividade inovativa – processos industriais complexos</b>	<b>Média atividade inovativa – dependentes de fornecedores</b>
Setores	Equipamentos de informática, produtos eletrônicos e ópticos-Produtos Farmoquímicos e farmacêuticos	Coque, derivados do petróleo e biocombustíveis; Fumo; Manutenção, máquinas e equipamentos; Metalurgia; Produtos de Madeira	Produtos químicos Máquinas, aparelhos e mat. elétricos; Máquinas e equipamentos; Veículos automotores; Produtos diversos; Couros e calçados	Impressão e reprodução; Produtos de metal; Alimentos; Bebidas; Têxtil; Vestuário; Celulose e papel; Borracha e plástico; Minerais não-metálicos; outros equipamentos de transporte; Móveis.
Taxa de inovação	Muito alta (82%), associadas com participação elevada de empresas com dispêndios em P&D (65%) e de cooperação para inovação (67,3%)	Baixa (média de 50%) inferiores à média da indústria de transformação, com participação reduzida de empresas com dispêndios em P&D (9,5%) de cooperação para inovação 22,1%)	Muito alta (82%), associada com participação média de empresas com dispêndios em P&D (48,6%) e de cooperação para inovação (41,3%)	Média-Baixa (média de 67%) inferiores à média da indústria de transformação, com baixa participação de empresas com dispêndios em P&D (29,2%) de cooperação para inovação (42%)
Inovação em produto	Alta (67,9%) com maior a participação de produtos novos para o mercado mundial (4,1%) entre os grupos	Muito baixa (20,9%) inferiores à média da indústria de transformação e foco em produtos novos para empresa (12,8%)	Alta (68,9%) com a maior participação de produtos novos para o mercado nacional (24,7%) entre os grupos	Média-Baixa (média de 47,5%) inferiores à média da indústria de transformação, com maior participação de produtos novos para empresa (33,9%)
Inovação em processos de negócios	Alta (73,9%) com maior importância relativa das funções de marketing, embalagem e pós-vendas (46,7%), gestão (44,9%) e processamento de informações e comunicações (39%);	Média-Baixa (média de 45,3%), com maior importância relativa das funções de gestão e organização do trabalho e RH (35,4%)	Alta (68,9%) com maior importância relativa das funções de gestão e organização do trabalho e RH (49%), gestão (46,6%) e produção (43,9%)	Média-Baixa (média de 54,7%) com maior importância relativa das funções de gestão e organização do trabalho e RH (38,2%), gestão (34,7%) e processamento de informações e comunicações (31,9%)

Fonte: Elaboração própria.

O Quadro 5 traz os 21 indicadores selecionados para os quatro agrupamentos identificados.

indicador/grupo	CLUSTER 0	CLUSTER 1	CLUSTER 2	CLUSTER 3	Total Indústria	Ind. transformação
TX-INOV	82,1%	53,2%	82,2%	68,9%	71,0%	71,4%
TX inov PROD	67,9%	20,9%	68,9%	47,5%	50,5%	51,3%
TX inov PROD-EMP	41,1%	12,8%	41,3%	33,9%	33,1%	34,0%
TX inov PROD-NAC	22,7%	6,6%	24,7%	10,7%	14,6%	14,6%
TX inov PROD-MUN	4,1%	1,5%	2,9%	2,9%	2,8%	2,8%
TX inov PROC	73,9%	45,3%	68,9%	54,7%	57,9%	58,3%
TX InovPROC-PROD	38,1%	21,4%	42,9%	28,2%	31,6%	31,9%
TX InovPROC-LOG	20,7%	14,3%	24,1%	18,7%	19,4%	19,8%
TX InovPROC-TIC	39,0%	20,9%	39,9%	31,9%	32,5%	33,2%
TX InovPROC-CONT	18,9%	21,5%	35,3%	25,5%	26,9%	27,6%
TX InovPROC-GEST	44,9%	28,6%	46,6%	34,7%	37,5%	37,7%
TX InovPROC-ORG	41,1%	35,4%	49,0%	38,2%	40,6%	41,0%
TX InovPROC-MKT	46,7%	14,6%	41,8%	33,5%	33,5%	34,3%
TX Inov-COOP	67,3%	22,1%	41,3%	42,0%	41,7%	41,3%
TX Inov-COOP-CLIEN	46,1%	17,3%	32,3%	31,5%	31,2%	31,2%
TX Inov-COOP-FORN	58,1%	13,0%	37,2%	35,6%	35,3%	35,2%
TX Inov-COOP-CONC	16,9%	4,2%	9,9%	7,9%	9,1%	8,6%
TX Inov-COOP-CT&I	48,1%	14,5%	26,0%	24,2%	25,7%	25,0%
TX Inov-COOP-STRUP	21,3%	9,8%	10,6%	5,7%	8,7%	8,2%
TX Inov-COOP-CONS	53,7%	14,1%	25,4%	28,6%	28,1%	27,5%
TX Inov-COOP-OEMP	37,6%	11,8%	15,6%	14,9%	16,1%	15,8%
TX Inov-DISP	65,0%	9,5%	48,6%	29,2%	33,9%	33,8%

**Quadro 5 - Indicadores selecionados da PINTEC Semestral 2021, por Cluster, Total da Indústria de Transformação e Total da Indústria, (em %)**

Fonte: Elaboração própria a partir da PINTEC Semestral, 2021.

#### 4.1. Taxas de inovação

A análise das empresas que implementaram inovações de produto e/ou processo de negócios, mostra que o Cluster 0 e Cluster 2 apresentam padrões semelhantes em termos da taxa geral de inovação e de inovação de produto. No Cluster 0, observa-se uma maior taxa de inovação em processos de negócios associada principalmente às atividades de marketing e gestão organizacional, enquanto o Cluster 2 apresenta taxas de inovação de produto ligeiramente superiores. O Cluster 1 é o que apresenta as meno-



res taxas em termos das empresas ativas em inovação, produto e processo, tanto em relação aos demais grupos como em relação à média da indústria de transformação. O Cluster 3 é aquele que apresenta as taxas de inovação geral, de produto e de processos de negócios mais próximas da média da indústria de transformação. A análise sobre o grau de novidade nas inovações de produto, por sua vez, evidencia que o cluster 0, composto pelos setores farmacêutico e de equipamentos de informática, eletrônica e óptica, apresentou a maior taxa de inovação em produtos novos para o mercado mundial (4,1%) associada principalmente ao desempenho do setor farmacêutico, onde 8,3% das empresas inovadoras declararam ter inovado em produtos para o mercado mundial, em comparação a um percentual de 0,6% do setor de equipamentos de informática, eletrônica e óptica. Nos Clusters 2 e 3, a taxa de inovação de produtos novos para o mercado mundial foi de 2,9%, muito próxima da média da indústria de transformação (2,8%), enquanto no Cluster 1, apenas 1,4% das empresas inovaram em produto para o mercado mundial. Quando observamos as inovações de produto novo para o mercado nacional, o grupo de setores do Cluster 2 foi o que apresentou o maior percentual (24,7%), seguido pelo Cluster 0 (22,7%). Este percentual é superior à média da indústria de transformação (14,6%), com médias muito inferiores à média da indústria de transformação.

A análise das inovações em processos de negócios mostra que os Clusters 0 e 2 também apresentaram as maiores taxas tanto em termos gerais como na maior parte das funções empresariais associadas a este tipo de inovação. O Cluster 0 apresentou uma taxa de inovação em processos de negócios de 73,9% enquanto no Cluster 2 esse percentual foi de 68,9%, caindo para 45,3% no cluster 1, enquanto no cluster 3 foi de 54,7%.

De um modo geral, a análise das taxas inovação em processo de negócios nas diferentes funções empresariais aponta para a maior importância relativa de funções associadas às práticas de gestão organizacional, marketing e organização do trabalho em todos os grupos ou Clusters. As funções de logística, entrega e distribuição e as funções de contabilidade e outras operações administrativas, por outro lado, foram as que apresentaram menor importância relativa entre as inovações de processos de negócios adotadas pelos diferentes grupos de setores. A baixa incidência de inovações de processos de negócios na função de logística, entrega ou distribuição, entre os grupos de setores pesquisados, parece se contrapor à percepção sobre a importância crescente que estas funções assumiram na economia, particularmente durante o período da pandemia da COVID-19. Entretanto,

é possível supor que inovações importantes no campo da logística são efetivamente adotadas em outros setores específicos de atividade econômica (seção H-CNAE 2.0 - transporte, armazenagem e correio).

#### 4.1.2. *Dispêndio em P&D*

A análise dos padrões de dispêndio em P&D na PINTEC Semestral 2021, conforme já destacado, se restringe a informações sobre o percentual de empresas que declararam ter realizado atividades de P&D em 2021. Neste aspecto, o Cluster 0 foi o que apresentou o maior percentual de empresas que realizaram dispêndio em atividades de P&D, com 65% das empresas em relação a 33,8% no total da indústria de transformação. Entretanto, cabe ressaltar que o número de empresas deste grupo corresponde 6,7% no universo total das empresas que declararam ter realizado dispêndios em P&D na indústria de transformação. O Cluster 2 também apresentou um percentual elevado de empresas que realizaram dispêndios em P&D no ano de 2021 (48,6%), além de contar com uma participação de 40,2% no total de empresas que declararam ter realizado dispêndios em P&D na indústria. Os Clusters 3 e 1 apresentaram percentuais de, respectivamente, 29,2% e 9,5% de empresas que declararam ter realizado dispêndios com P&D 2021. Dentre os setores que se destacam pelo elevado percentual de empresas inovadoras com dispêndio em P&D estão o de equipamentos de informática (66,6%), produtos químicos (65,3%), produtos farmacêuticos (63,0%) e máquinas e equipamentos (51,1%). A Tabela 1 mostra o percentual de empresas que realizaram atividades de P&D em 2021 por grupo de setores e a participação do grupo em relação ao total de 3.093 empresas que declararam ter realizado dispêndios com P&D 2021.

**Tabela 1 - Empresas que realizaram dispêndios em P&D e participação no total de empresas na indústria de transformação (%), 2021.**

Grupo	% empresas que realizaram P&D em 2021	Número de empresas	% participação do grupo no total de empresas que realizaram P&D
Cluster 0	65,0%	208	6,7%
Cluster 1	9,5%	93	3,0%
Cluster 2	48,6%	1243	40,2%
Cluster 3	29,2%	1549	50,1%
Indústria de transformação	33,8%	3093	100%

Fonte: Elaboração própria a partir de dados da PINTEC Semestral 2021.

### *4.1.3. Cooperação para inovação*

A análise dos padrões de cooperação na PINTEC procura evidenciar a participação ativa das empresas industriais em projetos conjuntos de P&D ou outros projetos de inovação com organizações que integram o sistema nacional de inovação. Os resultados da PINTEC Semestral 2021 mostram que 41,3% das empresas que inovaram em produto e/ou processo de negócios na indústria de transformação estabeleceram algum tipo de relação de cooperação com outras organizações em 2021. A análise dos padrões de cooperação a partir de grupos de setores mostra que as relações de cooperação para inovação tiveram uma frequência consideravelmente maior para o Cluster 0 (67,3%) do que para os demais grupos. No Cluster 3, 42,0% das empresas afirmaram ter inovado com relações de cooperação com outras organizações, enquanto nos Clusters 2 e 1 esse percentual foi de respectivamente 41,3% e 22,1%.

Em termos dos principais parceiros para inovação, as relações com fornecedores, clientes ou consumidores e consultores ou empresas de consultoria apresentaram os maiores percentuais de cooperação para o conjunto da indústria de transformação, respectivamente 35,2%, 31,2% e 27,5%. No caso do Cluster 0, observa-se a maior importância relativa das relações de cooperação para inovação com fornecedores (58,1%), consultores ou empresas de consultoria (53,7%) e com a infraestrutura de CT&I (48,1%). No Cluster 2, os três principais parceiros de cooperação para inovação em 2021 foram fornecedores (37,2%), clientes (31,8%) e infraestrutura de CT&I (25,8%). No caso do Cluster 3 observa-se um padrão de cooperação para inovação muito próximo do Cluster 2 com fornecedores, clientes e consultores como os três principais parceiros nas relações de cooperação para inovação.

No caso específico das relações de cooperação com startups, observa-se que 21,3% das empresas no Cluster 0 relataram haver inovado em cooperação com este tipo de parceiro. Nos demais grupos, entretanto, a frequência das relações de cooperação com startups foi consideravelmente menor, sendo de 10,6% no Cluster 2, 9,8% no Cluster 1 e 5,7% no Cluster 3. Adicionalmente, observa-se uma grande variação nas taxas de cooperação para inovação com startups entre os setores que integram cada grupo. Chama a atenção, por exemplo, a frequência elevada de relações de cooperação para inovação com startups na indústria extrativa (33,0%), nos setores de produtos diversos (14,3%), metalurgia (14,3%) e coque, derivados do petróleo e biocombustíveis (14,0%).

#### 4.2. *Aderência da agregação proposta com as taxonomias tecnológicas e setoriais tradicionais*

A classificação dos setores a partir destes quatro agrupamentos permite apontar especificidades nos padrões de esforço e desempenho inovativo dos setores da indústria de transformação brasileira, particularmente no tocante à sua aderência com as taxonomias de intensidade tecnológica e padrões setoriais já consagradas na literatura internacional. O Quadro 5 mostra uma comparação entre os clusters de setores obtidos através da metodologia adotada neste estudo, as classificações setoriais associadas à taxonomia de intensidade tecnológica da OCDE (2011) e à taxonomia de padrões setoriais de Pavitt (1984).

No caso do Cluster 0, que congrega setores intensivos em conhecimento, observa-se uma total convergência tanto em relação às taxonomias de intensidade tecnológica, como a de padrões setoriais de Pavitt. Cabe ressaltar, que a análise de grupos de setores a dois dígitos de Divisões CNAE, traz evidentes limitações analíticas na medida em que as taxonomias de intensidade tecnológica da OCDE e de padrões setoriais de Pavitt, contemplam diferenças nas classificações de grupos de setores (CNAE a 3 dígitos) que integram uma mesma divisão. No caso do setor de Fabricação de equipamentos de informática, produtos eletrônicos e ópticos (Divisão CNAE 26), observa-se que alguns grupos como a fabricação de componentes eletrônicos (Grupo CNAE 261) e a atividades de Fabricação de mídias virgens, magnéticas e ópticas (Grupo CNAE 268), são classificados como setores Difusores de Progresso Técnico na taxonomia de Pavitt, enquanto o setor Fabricação de equipamentos e instrumentos óticos, fotográficos e cinematográficos (Grupo CNAE 267) é classificado como setor Intensivo em Escala. Não obstante tais especificidades, observa-se a aderência das classificações tradicionais com os padrões de esforço e desempenho inovativo observados no Cluster 0.

O Cluster 1, que apresenta as menores taxas em termos dos indicadores de esforço e desempenho inovativo, apresenta algum grau de aderência com a taxonomia de intensidade tecnológica, na medida em que congrega dois setores com baixa intensidade (fumo e produtos de madeira), e três setores com intensidade tecnológica média-baixa (coque, derivados do petróleo e biocombustíveis; Metalurgia e Manutenção, reparação e instalação de máquinas e equipamentos). Entretanto, este agrupamento apresenta pouca aderência com a taxonomia de Pavitt, na medida em que agrupa

setores que integram três categorias distintas desta taxonomia (difusores de progresso técnico, intensivos em escala e dependentes de fornecedores), conforme mostra o Quadro 6.

O Cluster 2 congrega setores com processos industriais complexos e apresenta também uma maior aderência com a taxonomia de intensidade tecnológica. Este cluster congrega seis setores com taxas altas de esforço e desempenho inovativo de acordo com os indicadores construídos a partir da PINTEC Semestral 2021, sendo que quatro deles apresentam uma correspondência com classificação de média-alta intensidade tecnológica. Outros dois setores (calçados e produtos diversos), entretanto, são considerados como de baixa intensidade tecnológica de acordo com a taxonomia da OCDE. Com relação à taxonomia de padrões de concorrência de Pavitt, observa-se que dois dos setores que integram o Cluster 2 são considerados como difusores de progresso técnico (máquinas e equipamentos e máquinas, aparelhos e materiais elétricos), enquanto os demais se enquadram como intensivos em escala (veículos automotores, reboques e carrocerias e produtos químicos) ou dependentes de fornecedores (produtos diversos, couro e calçados).

O Cluster 3 é o que apresenta menor grau de aderência tanto com a taxonomia de intensidade tecnológica como com a de padrões de concorrência de Pavitt. Este cluster congrega 11 setores, dos quais sete são classificados como de baixa intensidade tecnológica, três como de média-baixa intensidade tecnológica e um como de média-alta intensidade tecnológica. Em termos da taxonomia de padrões de concorrência de Pavitt, esses 11 setores encontram-se distribuídos entre o grupo de setores intensivos em escala (cinco setores) e o grupo de setores dominados por fornecedores (seis setores).

Finalmente, cabe ressaltar, por um lado, que a própria natureza dos indicadores utilizados para agregação dos cluster de setores, associados a taxas de inovação, cooperação e dispêndio, tendem a aproximar essa classificação por clusterização da taxonomia de intensidade tecnológica. Por outro lado, as convergências e divergências observadas, na classificação dos Clusters de setores revelam particularidades importantes dos padrões de desempenho setorial na indústria de transformação brasileira, e que suscitam um aprofundamento da análise.

**Quadro 6 - Comparação entre as classificações de intensidade tecnológica da OCDE, padrões setoriais de Pavitt e análise de clusters da Pintec Semestral**

Taxonomia Intensidade Tecnológica	Análise de clusters	Taxonomia Pavitt
<b>Alta intensidade</b>	<b>CLUSTER 0</b>	<b>Grupo Baseados em Ciência</b>
Equipamentos de informática, produtos eletrônicos e ópticos	Produtos Farmoquímicos e farmacêuticos	Equipamentos de informática, produtos eletrônicos e ópticos
Produtos Farmoquímicos e farmacêuticos	Equipamentos de informática, produtos eletrônicos e ópticos	Produtos Farmoquímicos e farmacêuticos
<b>Média-alta Intensidade</b>	<b>CLUSTER 2</b>	<b>Grupo Difusores Prog Técnico</b>
Produtos químicos	Produtos químicos	Máquinas e equipamentos
Máquinas e equipamentos	Máquinas e equipamentos	Máquinas, aparelhos e materiais elétricos
Máquinas, aparelhos e materiais elétricos	Máquinas, aparelhos e materiais elétricos	Manutenção, reparação e instalação máquinas e equipamentos
Veículos automotores, reboques e carrocerias	Veículos automotores, reboques e carrocerias	<b>Grupo Intensivos em Escala</b>
Outros equipamentos de transporte	Produtos diversos	Bebidas
<b>Média-baixa Intensidade</b>	Couros e calçados	Veículos automotores, reboques e carrocerias
Produtos de metal	<b>CLUSTER 3</b>	Produtos alimentícios
Artigos de borracha e plástico	Impressão e reprodução de gravações	Metalurgia
Produtos de minerais não-metálicos	Produtos de metal	Produtos de metal
Coque, derivados do petróleo e biocombustíveis	Produtos alimentícios	Produtos de minerais não-metálicos
Metalurgia	Bebidas	Outros equipamentos de transporte
Manutenção, reparação e instalação máquinas e equipamentos	Produtos têxteis	Produtos químicos
<b>Baixa Intensidade</b>	Vestuário e acessórios	Coque, derivados do petróleo e biocombustíveis
Couros e calçados	Celulose, papel e produtos de papel	Fumo
Produtos diversos	Artigos de borracha e plástico	<b>Grupo Dependentes de Fornecedores</b>
Celulose, papel e produtos de papel	Produtos de minerais não-metálicos	Produtos diversos
Produtos alimentícios	Outros equipamentos de transporte	Produtos têxteis
Bebidas	Móveis	Couros e calçados
Fumo	<b>CLUSTER 1</b>	Móveis
Impressão e reprodução de gravações	Fumo	Celulose, papel e produtos de papel
Produtos têxteis	Produtos de Madeira	Vestuário e acessórios
Móveis	Coque, derivados do petróleo e biocombustíveis	Impressão e reprodução de gravações
Vestuário e acessórios	Metalurgia	Artigos de borracha e plástico
Produtos de Madeira	Manutenção, reparação e instalação máquinas e equipamentos	Produtos de Madeira

Fonte: Elaboração própria.

## 5. Considerações Finais

Alguns princípios gerais condicionam a “funcionalidade” ou a “eficácia” de uma determinada taxonomia tecnológica setorial. Em primeiro lugar, a taxonomia deve ter um nível suficiente de “granularidade” para refletir com precisão a diversidade dos setores, evitando a generalização excessiva ou a complexidade excessiva. Em segundo lugar, a taxonomia deve fornecer critérios claros para categorizar setores, com base em fatores observáveis e mensuráveis. Em terceiro lugar, a taxonomia deve incluir diferentes tipos de inovação, especialmente para setores onde a inovação não tecnológica desempenha um papel importante. Em quarto lugar, a taxonomia deve refletir as interdependências entre os setores, reconhecendo que a inovação em um setor (por exemplo, TIC) pode ter efeitos colaterais em outros setores (por exemplo, saúde ou logística). Em quinto lugar, a taxonomia deve ser empiricamente sólida, estando baseada em dados robustos e pesquisas empíricas estruturadas, garantindo que as classificações sejam baseadas no desempenho real da inovação setorial. Em sexto lugar, taxonomias setoriais devem ser flexíveis o suficiente para levar em conta mudanças ao longo do tempo, como rupturas tecnológicas ou convergência setorial, devendo ser capazes de evoluir com novos dados, tecnologias e mudanças nos limites da indústria. Em sétimo lugar, a taxonomia deve ser amigável ao usuário e acessível a uma ampla gama de partes interessadas, incluindo formuladores de políticas, pesquisadores e profissionais da indústria, apresentando informações de uma forma que seja compreensível e prática para diferentes usuários. Por fim, a taxonomia deve levar a recomendações acionáveis tanto para formuladores de políticas quanto para líderes empresariais, ajudando-os a elaborar políticas, alocar recursos ou investir em setores específicos.

Considerando estes princípios e buscando captar particularidades dos padrões de esforços e desempenho inovativo na indústria de transformação brasileira, o artigo propôs a construção de uma metodologia para agregação setorial a partir dos dados da PINTEC Semestral 2021. O exercício de agregação proposto permitiu a identificação de quatro agrupamentos de setores com características distintas de acordo com um conjunto de 21 variáveis relacionadas com o desempenho e esforço inovativo das empresas. Em especial, os resultados apontaram para algumas especificidades nos padrões de inovação dos setores da indústria de transformação brasileira no tocante à sua aderência com as taxonomias de intensidade tecnológica e padrões setoriais já consagradas na literatura internacional.

No âmbito dos agrupamentos identificados, destaca-se inicialmente um conjunto restrito composto por dois setores intensivos em conhecimento (Cluster 0), que além de apresentar taxas de inovação elevadas responde por um percentual elevado de empresas inovadoras que realizam dispêndios em P&D interno. Esse grupo de setores se destaca também pela intensidade das relações de cooperação com outros parceiros para inovação, particularmente com organizações ligadas à infraestrutura de CT&I.

Em termos de dinamismo inovativo, esse primeiro agrupamento de setores se assemelha ao Cluster 2, composto majoritariamente por setores que operam com processos industriais complexos. Entretanto, ainda que o agrupamento ligado ao Cluster 2 apresente taxas de inovação (geral de produto e de processos de negócios) muito próximas daquelas observadas no primeiro agrupamento mencionado (Cluster 0), ele se diferencia pelas menores taxas de cooperação para inovação e pela menor participação relativa de empresas com dispêndio em atividades de P&D.

No extremo oposto, foi possível identificar dois outros agrupamentos de setores (Cluster 1 e Cluster 3), com menor desempenho e esforço inovativo, compostos majoritariamente por setores intensivos em escala ou dependentes de fornecedores. O Cluster 1 se caracteriza por apresentar as menores taxas de inovação, inferiores à média da indústria de transformação, com participação muito reduzida de empresas inovadoras que realizaram dispêndios em P&D e reduzido percentual de empresas que inovaram através do estabelecimento de relações de cooperação com outros parceiros. Já o Cluster 3 é aquele que congrega o maior número de setores da indústria de transformação (11 setores) e apresenta maior heterogeneidade quando analisados à luz de taxonomias de intensidade tecnológica ou padrões de concorrência.

Cabe ressaltar que o exercício realizado foi de natureza essencialmente exploratória, visando identificar padrões de realização de atividades inovativas no plano setorial, informados pelas classificações internacionais, mas que considerem especificidades da realidade da indústria brasileira. Esse esforço pode ter desdobramentos normativos importantes na identificação de padrões que possibilitem ajustamentos posteriores no instrumento de coleta de informações da PINTEC-Semestral. Além disso, a análise realizada pode ser refinada a partir de associações dos padrões identificados a estímulos, benefícios e impactos das atividades inovativas realizadas, captados a partir de informações qualitativas levantadas. O mesmo se aplica



em relação à identificação de possíveis influências de políticas públicas sobre aqueles padrões, aspecto incorporado de forma mais detalhada nas edições seguintes da PINTEC-Semestral.

Algumas limitações e possíveis desdobramentos da análise podem ser destacados. A análise de grupos de setores a dois dígitos (nível de Divisões da CNAE), trouxe limitações analíticas para o exercício metodológico proposto. Entretanto, os padrões observados e os resultados obtidos no exercício de agrupamento de setores permitiram evidenciar particularidades importantes dos padrões de desempenho inovativo na indústria de transformação brasileira, que suscitam um aprofundamento da análise a partir de uma base de dados com maior nível de desagregação setorial e acesso aos microdados ao nível da firma.

A possibilidade de validações adicionais a partir de novas rodadas da PINTEC Semestral é um caminho natural, que deverá ser explorado em novas análises, e, nesse sentido, a análise elaborada no artigo pode ser vista como um ponto inicial de referência, a partir do qual padrões setoriais mais dinâmicos podem ser vislumbrados, eventualmente gerando processos simultâneos de reagrupamento de setores. A ferramenta metodológica utilizada pode ser particularmente útil para avançar nessa direção, representando uma evolução natural nas metodologias tradicionais de análise de clusters, ajustadas ao contexto atual de Big Data e análise computacional, através do aprendizado de máquina não supervisionado, o qual representa uma evolução da ciência de dados, introduzindo ferramentas da ciência da computação ao campo da Análise Multivariada.

Por fim, cabe ressaltar que, no contexto atual marcado pela incorporação das novas plataformas tecnológicas digitais associadas à chamada Indústria 4.0 e pela crescente servitização das atividades industriais, os limites fundamentais que determinam as relações e interações intersetoriais estão se desfazendo e se dissolvendo rapidamente. Cada vez mais, empresas e setores operam em sistemas de inovação complexos, dinâmicos e adaptativos onde as fronteiras entre produtos e serviços estão se desfazendo e reduzindo as linhas de demarcação tradicionais propostas pelas classificações setoriais baseadas nos processos de produção. Tal fenômeno, entre outras implicações, coloca em xeque a própria análise de padrões tecnológicos e de concorrência a partir das classificações setoriais tradicionais e deve pautar uma nova frente de desafios metodológicos para as pesquisas de inovação.

**ANEXO1 - Divisões da Classificação Nacional de Atividades Econômicas – CNAE 2.0 das atividades da Indústria de transformação e Indicadores da PINTEC Semestral 2021**

Sector	TX+ NOV	TX inv PROD	TX inv PROD- EMP	TX inv PROD- NAC	TX inv PROD- MUN	TX inv PROD- PROD	InovPROC- LOG	InovPROC- C TIC	InovPROC- CONT	InovPROC- GEST	InovPROC- ORG	InovPROC- MKT	Inov- COOP- COOP	Inov- COOP- CLIE	Inov- COOP- FORN	Inov- COOP- CONC	Inov- COOP- CT&I	Inov- COOP- STRUP	Inov- COOP- CONS	Inov- COOP- DEMP	Inov- DISP
Artigos de borracha e plástico	0,73	0,62	0,50	0,12	0,00	0,44	0,19	0,09	0,28	0,23	0,24	0,32	0,24	0,40	0,39	0,05	0,16	0,07	0,23	0,15	0,35
Bebidas	0,77	0,66	0,59	0,07	0,00	0,53	0,37	0,20	0,17	0,12	0,47	0,35	0,26	0,53	0,34	0,53	0,00	0,44	0,05	0,46	0,32
Celulose, papel e produtos de papel	0,72	0,44	0,28	0,15	0,01	0,63	0,32	0,06	0,45	0,15	0,51	0,47	0,43	0,37	0,35	0,36	0,13	0,19	0,02	0,21	0,37
Coque, derivados do petróleo e biocombustíveis	0,59	0,18	0,12	0,05	0,01	0,45	0,07	0,05	0,24	0,20	0,14	0,33	0,13	0,18	0,11	0,14	0,02	0,15	0,14	0,15	0,17
Couros e calçados	0,68	0,51	0,42	0,09	0,01	0,59	0,44	0,42	0,50	0,41	0,49	0,56	0,37	0,39	0,38	0,39	0,14	0,26	0,01	0,14	0,43
Equipamentos de informática, produtos eletrônicos e ópticos	0,87	0,66	0,38	0,28	0,01	0,79	0,41	0,25	0,35	0,14	0,43	0,36	0,50	0,67	0,46	0,59	0,18	0,52	0,22	0,52	0,67
Fumo	0,64	0,25	0,09	0,00	0,16	0,57	0,39	0,27	0,18	0,09	0,35	0,18	0,26	0,06	0,00	0,06	0,00	0,06	0,06	0,06	0,29
Impressão e reprodução de gravações	0,75	0,51	0,33	0,00	0,18	0,62	0,26	0,15	0,28	0,24	0,35	0,35	0,59	0,16	0,00	0,16	0,16	0,16	0,00	0,00	0,27
Manutenção, reparação e instalação maq equip	0,52	0,22	0,20	0,02	0,00	0,43	0,30	0,08	0,22	0,31	0,31	0,42	0,03	0,19	0,17	0,01	0,02	0,02	0,01	0,01	0,01
Máquinas e equipamentos	0,83	0,72	0,48	0,18	0,05	0,71	0,33	0,18	0,40	0,36	0,51	0,56	0,42	0,45	0,41	0,45	0,08	0,30	0,14	0,28	0,51
Máquinas, aparelhos e materiais elétricos	0,78	0,64	0,34	0,30	0,00	0,62	0,34	0,18	0,46	0,45	0,40	0,39	0,48	0,42	0,32	0,32	0,19	0,13	0,12	0,31	0,41
Metallurgia	0,58	0,23	0,09	0,11	0,02	0,49	0,22	0,21	0,30	0,20	0,45	0,32	0,29	0,28	0,26	0,17	0,04	0,27	0,14	0,25	0,16
Móveis	0,71	0,53	0,38	0,15	0,00	0,51	0,26	0,10	0,25	0,10	0,18	0,29	0,28	0,35	0,24	0,28	0,01	0,22	0,06	0,23	0,23
Outros equipamentos de transporte	0,71	0,36	0,27	0,07	0,02	0,60	0,33	0,07	0,24	0,44	0,46	0,56	0,37	0,27	0,26	0,27	0,01	0,25	0,04	0,22	0,08
Produtos alimentícios	0,81	0,40	0,33	0,04	0,03	0,52	0,22	0,25	0,33	0,28	0,35	0,36	0,35	0,49	0,34	0,40	0,12	0,29	0,04	0,43	0,34
Produtos de Madeira	0,43	0,19	0,11	0,09	0,00	0,43	0,20	0,19	0,06	0,16	0,16	0,36	0,09	0,23	0,11	0,23	0,11	0,11	0,11	0,13	0,02
Produtos de metal	0,77	0,53	0,19	0,26	0,08	0,67	0,45	0,19	0,23	0,19	0,35	0,39	0,36	0,39	0,33	0,26	0,07	0,25	0,08	0,18	0,41
Produtos de minerais não-metálicos	0,68	0,46	0,37	0,05	0,03	0,59	0,36	0,28	0,39	0,39	0,44	0,55	0,54	0,44	0,33	0,35	0,08	0,25	0,12	0,34	0,24
Produtos diversos	0,82	0,74	0,50	0,21	0,04	0,71	0,48	0,18	0,29	0,30	0,50	0,47	0,41	0,46	0,28	0,41	0,03	0,30	0,14	0,41	0,38
Produtos Farmoquímicos e farmacêuticos	0,77	0,70	0,45	0,17	0,08	0,68	0,34	0,16	0,44	0,25	0,47	0,48	0,42	0,68	0,46	0,56	0,15	0,43	0,20	0,56	0,63
Produtos químicos	0,90	0,76	0,36	0,36	0,05	0,74	0,41	0,19	0,40	0,29	0,36	0,42	0,52	0,30	0,26	0,25	0,01	0,20	0,13	0,21	0,65
Produtos têxteis	0,81	0,57	0,41	0,12	0,05	0,46	0,35	0,07	0,36	0,26	0,27	0,22	0,27	0,43	0,30	0,32	0,06	0,26	0,07	0,19	0,14
Veículos automotores, rebocadores e carrocerias	0,85	0,70	0,38	0,31	0,00	0,71	0,59	0,32	0,34	0,34	0,53	0,49	0,30	0,49	0,28	0,42	0,18	0,32	0,07	0,23	0,40
Vestilário e acessórios	0,63	0,36	0,21	0,14	0,01	0,62	0,24	0,29	0,38	0,32	0,41	0,52	0,23	0,38	0,21	0,38	0,08	0,18	0,01	0,25	0,08

Fonte: Elaboração própria a partir de dados da PINTEC Semestral 2021.

## Bibliografia

- Alpaydin, Ethem. 2010. *Introduction to Machine Learning*. 3rd ed. Massachusetts: MIT Press.
- Archibugi, Daniele. 2001. "Pavitt's Taxonomy Sixteen Years On: A Review Article." *Economics of Innovation and New Technology* 10: 415–425.
- Arocena, Rodrigo, e Judith Sutz. 2000. "Looking at National Systems of Innovation from the South." *Industry and Innovation* 7: 55–75.
- Campos, B., e Ana Urraca-Ruiz. 2009. "Padrões Setoriais de Inovação na Indústria Brasileira." *Revista Brasileira de Inovação* 8 (1): 167–210.
- Castellaci, Fulvio. 2008. "Technological Paradigms, Regimes and Trajectories: Manufacturing and Service Industries in a New Taxonomy of Sectoral Patterns of Innovation." *Research Policy* 37 (6–7): 978–994.
- Castellaci, Fulvio. 2009. "The Interactions between National Systems and Sectoral Patterns of Innovation: A Cross-Country Analysis of Pavitt's Taxonomy." *Journal of Evolutionary Economics* 19: 321–347. <https://doi.org/10.1007/s00191-008-0113-9>.
- Castro, Daniela Fernandes de. 2010. *Padrões Setoriais da Inovação Tecnológica na Indústria Brasileira: Uma Análise de Cluster a Partir da PINTEC*. Tese de Doutorado, Pontifícia Universidade Católica de São Paulo, São Paulo.
- Cavalcante, L. R. 2014. *Classificações Tecnológicas: Uma Sistematização*. Brasília: Instituto de Pesquisa Econômica Aplicada (Ipea). Nota Técnica Diset, n. 17.
- Cimoli, Mario, e Gabriel Porcile. 2011. "Learning, Technological Capabilities, and Structural Dynamics." In *The Oxford Handbook of Latin American Economics*. New York: Oxford University Press.
- Dangeti, Pratap. 2017. *Statistics for Machine Learning*. Birmingham, UK: Packt Publishing.
- Edquist, Charles. 1997. *Systems of Innovation: Technologies, Institutions and Organisations*. London: Pinter.
- European Commission. 2008. *Sectoral Innovation Watch: Synthesis Report*. Europe INNOVA Paper N° 8.
- Fávero, Luiz Paulo. 2022. *Análise de Dados: Estatística e Modelagem Multivariada com Excel, Stata, SPSS e Python*. Rio de Janeiro: LTC.
- Ferraz, João Carlos, e Jorge Nogueira Britto. 2020. *Projeto Indicadores de Inovação e Digitalização (PIID): Avaliações e Sugestões para Medições Efetivas e Eficazes*. Relatório Final. Confederação Nacional da Indústria (CNI), setembro.
- Ferraz, João Carlos, Jorge Nogueira Britto, Mariana Szapiro, F. V. Silva, A. O. M. Pinheiro, F. J. M. Peixoto, A. V. Rodrigues, e M. F. Ribeiro. 2022. "Short Period Innovation Survey: General Guidelines and the Brazilian Experience." *26th International Conference on Science, Technology, and Innovation Indicators (STI)*, 7–9 September, Granada, Spain.
- Furtado, André Tosi e Ruy de Quadros Carvalho. 2005. "Padrões de Intensidade Tecnológica da Indústria Brasileira: Um Estudo Comparativo com os Países Centrais." *São Paulo em Perspectiva* 19 (1): 70–84.
- Gault, Fred. 2018. "Defining and Measuring Innovation in All Sectors of the Economy." *Research Policy* 47 (3): 617–622.
- Guido, Sarah, e Andreas Muller. 2016. *Introduction to Machine Learning with Python*. [S.l.]: O'Reilly Media.
- Hartigan, J. A., e M. A. Wong. 1979. "Algorithm AS 136: A K-means Clustering Algorithm." *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 28 (1): 100–108. <https://doi.org/10.2307/2346830>. Acesso em 21 mar. 2023.
- Hermida, Camila, e Célio Lourenço Xavier. 2012. "Competitividade da Indústria Brasileira no Período Recente de Acordo com a Taxonomia de Pavitt." *Revista Brasileira de Inovação* 11 (2): 365–396.
- Hatzichronoglou, Thomas. 1997. "Revision of the High-Technology Sector and Product Classification." Working Paper no. 1997/2. Paris: OECD Publishing.

- Hinloopen, Jeroen. 2003. "Innovation Performance across Europe." *Economics of Innovation & New Technology* 12 (2): 145–161.
- Instituto Brasileiro de Geografia e Estatística (IBGE). 2022. *Pesquisa de Inovação Semestral: 2021: Indicadores Básicos / IBGE, Coordenação de Estatísticas Estruturais e Temáticas em Empresas*. Rio de Janeiro: IBGE.
- Jain, Anil K. 2010. "Data Clustering: 50 Years beyond K-means." *Pattern Recognition Letters* 31 (8): 651–666. <https://doi.org/10.1016/j.patrec.2009.09.011>. Acesso em 21 mar. 2023.
- Kannebley, Sergio, Geciane Porto, e Elaine Toldo Pazello. 2005. "Characteristics of Brazilian Innovative Firms: An Empirical Analysis Based on PINTEC—Industrial Research on Technological Innovation." *Research Policy* 34 (6).
- Laplane, M. F., F. Sarti, R. Sabbatini, e C. Hiratuka. 2001. "El Caso Brasileño." In *El Boom de Inversión Extranjera Directa en el Mercosur*, editado por Daniel Chudnovsky, 123–208. Buenos Aires: Sigloveintiuno.
- Lucena, Rodrigo Milano de. 2022. "Especificidades no Padrão Setorial de Inovação em Países em Desenvolvimento: Uma Investigação para a Indústria Brasileira." *Textos de Economia* 25 (2): 1–22.
- Malerba, Franco. 1992. "Learning by Firms and Incremental Technical Change." *The Economic Journal* 102: 845–859.
- Malerba, Franco. 2002. "Sectoral Systems of Innovation and Production." *Research Policy* 31: 247–264.
- Malerba, Franco. 2007. "Innovation and the Dynamics and Evolution of Industries: Progress and Challenges." *International Journal of Industrial Organization* 25: 675–699.
- Malerba, Franco., e Luigi Orsenigo. 1997. "Technological Regimes and Sectoral Patterns of Innovative Activities." *Industrial and Corporate Change* 6: 83–117.
- Marques, Mabel Diz, José Eduardo Roselino, e Suelene Mascarini. 2019. "Taxonomias Tecnológicas e Setoriais da Indústria de Transformação Brasileira." *Revista Brasileira de Inovação* 18 (2): 417–448.
- Moura, F., Carina Nakatani-Macedo, M. Camara, e K. Maia. 2015. "Comportamento da Indústria Nacional por Segmentos de Intensidade Tecnológica em 2011: Uma Proposta Metodológica de Análise." *Revista Brasileira de Economia de Empresas* 15: 25–48.
- Morettin, Pedro Alberto, e Júlio da Motta Singer. 2023. *Estatística e Ciência de Dados*. Rio de Janeiro: LTC.
- OECD/Eurostat. 2018. *Oslo Manual 2018: Guidelines for Collecting, Reporting and Using Data on Innovation*. 4th ed. The Measurement of Scientific, Technological and Innovation Activities. Paris: OECD Publishing. <https://doi.org/10.1787/9789264304604-en>.
- OECD. 2011. *ISIC Rev. 3 Technology Intensity Definition*. OECD Directorate for Science, Technology and Industry, julho.
- OECD. 2007. *OECD Science, Technology and Industry Scoreboard 2007*. Annex 1: "Classification of Manufacturing Industries Based on Technology," 219–221. Paris: OECD.
- Pavitt, Keith 1984. "Sectoral Patterns of Technical Change: Towards a Taxonomy and a Theory." *Research Policy* 13: 343–373.
- Peneder, Michael 2007. "A Sectoral Taxonomy of Educational Intensity." *Empirica* 34 (3): 189–212.
- Rezende, Solange Oliveira. 2005. *Sistemas Inteligentes: Fundamentos e Aplicações*. [S.l.]: Manole.
- Robson, M., J. Townsend, e K. Pavitt. 1988. "Sectoral Patterns of Production and Use of Innovations in the UK: 1945–1983." *Research Policy* 17 (1): 1–14.
- Silva, C. F., e W. Suzigan. 2014. "Padrões Setoriais de Inovação da Indústria de Transformação Brasileira." *Estudos Econômicos* 44 (2): 277–321.
- Tan, Pang-Ning, Michael Steinbach e Vipin Kumar. 2014. *Introduction to Data Mining*. [S.l.]: Pearson.
- Townsend, J., F. Henwood, G. Thomas, K. Pavitt, e S. Wyatt. 1981. *Innovations in Britain since 1945*. Occasional Paper n. 16. Science Policy Research Unit, University of Sussex.

Yonamini, Fernanda Marie. 2011. *Nova Taxonomia de Regimes Tecnológicos para o Caso de um País em Desenvolvimento como o Brasil*. Tese (Doutorado em Desenvolvimento Econômico), Universidade Federal do Paraná, Curitiba.

## DECLARAÇÃO DE DISPONIBILIDADE DE DADOS

Os dados utilizados neste estudo estão disponíveis mediante solicitação ao autor. Dados adicionais e informações complementares também poderão ser fornecidos para fins de verificação ou replicação. A disponibilização está condicionada à inexistência de restrições de acesso público.

## CONTRIBUIÇÕES DE AUTORIA

**MV:** Conceitualização, Curadoria de dados, Análise formal, Investigação, Metodologia, Visualização, Escrita - rascunho original e Escrita - revisão e edição.

**RM:** Conceitualização, Curadoria de dados, Análise formal, Investigação, Metodologia, Visualização, Escrita - rascunho original e Escrita - revisão e edição.

**JB:** Conceitualização, Curadoria de dados, Análise formal, Investigação, Escrita - rascunho original e Escrita - revisão e edição.

**JF:** Conceitualização, Curadoria de dados, Análise formal, Investigação e Escrita - revisão e edição.

**MS:** Conceitualização, Curadoria de dados, Análise formal, Investigação e Escrita - revisão e edição.

## CONFLITO DE INTERESSE

Os autores declaram não terem quaisquer conflitos de interesse.

## EDITOR-CHEFE

Dante Mendes Aldrighi  <https://orcid.org/0000-0003-2285-5694>

Professor - Department of Economics University of São Paulo (USP)