

ANEXO I – DIRETRIZES TÉCNICAS PARA ELABORAÇÃO DOS ESTUDOS

Os dados devem ser fornecidos em Base de Dados contendo os atributos mínimos definidos neste Anexo.

1. RELATÓRIO DE METODOLOGIA

Deverá ser elaborado, após reunião inicial para fins de alinhamento a ser realizada entre contratada e contratante no prazo máximo de até dez dias após a celebração do contrato, um relatório discorrendo sobre as etapas, as análises e as inferências a serem utilizadas no desenvolvimento da Base de Dados e dos métodos de geração dos dados.

2. BASE DE DADOS

A Base de Dados deverá ser desenvolvida conforme atributos mínimos dos dados a serem entregues conforme descrito nesse Termo de Referência e acordado na Reunião de Kick-Off, e contemplará, no mínimo, as seguintes especificações:

Origem e natureza dos dados

Os dados devem ser agregados, expressos em número de viagens por fluxo mensal entre e os municípios, contendo os atributos mínimos apresentados nesse Termo de Referência e serem provenientes de:

- 1) registros dos usuários de telefonia móvel; e/ou
- 2) registros de GPS; e/ou
- 3) outras potenciais fontes de dados acerca do deslocamento de indivíduos entre e os municípios brasileiros.

Formato dos dados

Os dados devem ser fornecidos em Base de Dados contendo os atributos mínimos definidos no Projeto Básico.

Os dados devem ser anonimizados, e os fluxos devem contemplar os deslocamentos entre e os 5570 municípios brasileiros, independentemente do modo de transporte utilizado.

Para os deslocamentos entre municípios realizados pelo modo aéreo, a ser identificado por meio da análise do tempo de deslocamento entre municípios, devem ser informados os municípios de origem e/ou destino reais do passageiro, além dos aeródromos de origem e/ou de destino conforme cadastro de aeródromos públicos ou privados da ANAC.

Durante o processamento dos dados, quando identificado um novo aeródromo que não estiver presente no cadastro de aeródromos públicos ou privados, aquele deverá ser incluído como local de pouso e decolagem identificado. Para a realização da inclusão do novo aeródromo, deverão ser seguidas duas observações: a criação de um identificador próprio para aquele ponto no município, seguida da inclusão da sua coordenada geográfica.

A classificação do modo de transporte dos deslocamentos deverá ser feita, no mínimo, entre as classes: deslocamento aéreo de pessoas; deslocamento não aéreo de pessoas particular (que compreende o transporte terrestre ou aquaviário); deslocamento não aéreo coletivo de pessoas (que compreende o transporte terrestre ou aquaviário) e; potencial deslocamento para transporte de cargas. Essa classificação deverá ser feita com base nos atributos e em análises avançadas de aprendizado de máquina do ramo da inteligência artificial, buscando identificação e extrapolação de padrões de viagens de cada grupo, considerando, no mínimo: o tempo de deslocamento do usuário no par O/D; a frequência do viajante; os potenciais pontos de parada que podem indicar a natureza da viagem; a presença de grandes grupos de pessoas em rotas similares, que podem indicar a característica de transporte coletivo e; a disponibilidade de infraestrutura de transportes. Para isso, podem ser utilizados dados amostrais proporcionalmente suficientes provenientes da CONTRATANTE, que possam ser cruzados com os dados da CONTRATADA.

Serão caracterizados como origem ou destino os locais (municípios) em que o registro permanecer estacionado por período superior a 6 (seis) horas, conforme Wang (2013), com casos particulares ou exceções em acordo com a CONTRATANTE.

Os dados também devem ser classificados em relação à frequência dos viajantes no período, e em relação à velocidade média calculada, com classes a serem acordadas entre CONTRATANTE e CONTRATADA.

Escopo temporal

Os dados devem abranger o período de 24 (vinte e quatro) meses e devem ser coletados de forma ininterrupta a partir do ano de 2018. O primeiro período específico de doze meses será estabelecido na reunião inicial para fins de alinhamento a ser realizada entre contratada e contratante no prazo máximo de até dez dias após a celebração do contrato.

Aspectos relacionados à qualidade dos dados

A contratada deve comprovar significância estatística de acordo com metodologia renomada sobre o tema. Da mesma forma, deve apresentar amostra mínima e nível de confiabilidade em relação ao universo e a abrangência geográfica condizente ao escopo do trabalho, sendo necessária a cobertura mínima de 81% dos municípios. Na comprovação de amostra mínima e nível de confiabilidade, deve ser respeitada a abrangência geográfica em todos os Estados e Regiões.

De acordo com a Anatel (2016), para os serviços de telefonia móvel, deve-se considerar como cobertura o mínimo de 80% da área urbana das sedes dos municípios, sendo seu cumprimento acompanhado pela fiscalização da Anatel periodicamente. Levando este parâmetro em consideração, a contratada deve comprovar ter dados referentes à cobertura, no mínimo, do padrão 3G em pelo menos 80% de cada uma das 137 mesorregiões brasileiras como definidas pelo Instituto de Pesquisa Econômica Aplicada (IPEA, 2014).

3. METADADOS

Deverá ser disponibilizada juntamente com a Base de Dados, a documentação dos Metadados.

Entende-se por metadado os dados que fornecem informações sobre outros dados. De acordo com a Secretaria de Tecnologia da Informação, órgão federal vinculado ao Ministério do Planejamento, Orçamento e Gestão “[...] metadados são utilizados para registrar atributos sobre um recurso informacional visando facilitar a recuperação, a gestão, a interoperabilidade, dar suporte à identificação digital e dar suporte ao arquivamento e preservação” (BRASIL, 2017). Ademais, através dos Metadados é possível obter um conhecimento preliminar das propriedades e das características de um dado. Uma Base de Dados que possua uma documentação através dos seus Metadados terá seu conteúdo contextualizado, e também será possível, por exemplo, a própria identificação da autenticidade e da qualidade dos dados.

Além dos dados e informações que contextualizem a Base de Dados, deverão ser abordadas na documentação dos Metadados as seguintes informações relativas a aspectos qualitativos e à manipulação da Base de Dados:

- nível de tecnologia de cobertura da rede (2G, 3G, 4G, 5G) e sua partição;
- outras tecnologias consideradas no levantamento, entre elas, Wi-Fi, Bluetooth, GPS;
- especificação da utilização de Registros de Dados de Chamadas (CDRs) somente ou outros (citando os registros);
- especificação da utilização de dados ativos somente ou de dados passivos também;
- especificação das fontes dos dados utilizados (mais de um operador de telefonia móvel/outras fontes) para a composição da base final;
- métodos de validação aplicados;
- se os dados correspondem a dados já coletados/manipulados ou se foram coletados/manipulados para atender ao objeto da contratação.

4. TABELA DE DADOS AUXILIARES

Deverá ser disponibilizado juntamente com a Base de Dados, a Tabela de Dados Auxiliares.

As tabelas de dados auxiliares contêm o detalhamento de um item participante de uma determinada relação com a tabela principal. Estas tabelas são comumente utilizadas em modelos entidade-relacionamento em banco de dados. O relacionamento ocorre através da relação das chaves primárias das tabelas envolvidas.

5. RELATÓRIO DESCRITIVO DO MÉTODO DE GERAÇÃO DA BASE DE DADOS;

Deverá ser disponibilizado juntamente com a Base de Dados um arquivo contendo relatório que descreva, de forma sucinta, as informações agregadas à Base de Dados, a fim de validar os passos e as inferências apresentadas no Relatório de Metodologia, além de citar outros eventos relevantes que possam vir a ocorrer.

O Relatório Descritivo do Método de Geração da Base de Dados deverá delinear:

- tipos de eventos considerados válidos para geração da Base de Dados;

- cobertura espacial, período de abrangência e justificativas de quaisquer eventos desprezados;
- se foram assinaladas tendências de deslocamento e de que modo foram levantadas;
- se dados secundários foram utilizados e de que forma;
- detalhamento de qualquer outra manipulação que tenha sido realizada nos dados;
- se há restrição de licença em torno do uso do produto (dados) para uma finalidade específica e qual o potencial de reutilização de dados para outros fins, estudos ou pesquisas futuras.

6. SUPORTE TÉCNICO ESPECIALIZADO

O Suporte Técnico especializado, que deverá ser disponibilizado pelo período de 6 (seis) meses após a disponibilização da Base de Dados, tem por escopo o esclarecimento de dúvidas que venham a surgir durante o uso dos dados pela contratante relativas à transferência e à importação dos dados, o suporte na validação dos dados, bem como a garantia da qualidade dos dados, ou seja, caso seja verificada alguma inconsistência, é necessário que a contratada dê todo o suporte necessário para que este problema seja resolvido, até mesmo via atualização dos dados, caso necessário.

7. CRITÉRIOS DE ACEITAÇÃO DO OBJETO

Deverá ser efetuada, entre a contratada e a contratante, em até dez dias após a emissão do Pedido de Compra/Serviço, uma reunião de alinhamento para fins de detalhamento da proposta metodológica a ser utilizada no Relatório de Metodologia. Esta reunião de alinhamento deverá ocorrer em Brasília-DF em local, data e hora a serem definidos pela contratante, ou por meio remoto.

A entrega do Relatório de Metodologia, da Base de Dados, dos Metadados, da Tabela de Dados auxiliares e do Relatório Descritivo do Método de Geração da Base de Dados, apriori, deverá ser por meio de um File Transfer Protocol (FTP) privado, a ser disponibilizado pela contratada, cujos acessos deverão ser controlados por intermédio de login e senha. Todos os acessos ao FTP devem ser disponibilizados juntamente com a entrega do Relatório de Metodologia, devendo permanecer disponíveis durante toda a vigência do contrato. O meio de disponibilização poderá ser alterado caso as partes acordem nova alternativa na reunião inicial de alinhamento.

O sistema de arquivos a ser utilizado para disponibilização dos dados poderá ser, por exemplo, .csv, backup de SGBD, webservices. Todavia, deverá ser estabelecida extensão final entre as partes durante a reunião de alinhamento.

A Base de Dados poderá ser particionada em diferentes arquivos, em eventual restrição de tamanho físico, desde que acordado previamente com a contratante. Entretanto, quando disponibilizado de forma particionada, esta deverá estar acompanhada de documentação para orientar a junção e a identificação da sua ordem.

A contratada deverá prover Suporte Técnico Especializado a ser prestado por meio de e-mail, telefone, ou videoconferência durante o período de 6 (seis) meses após a disponibilização da Base de Dados. O Suporte Técnico Especializado engloba o auxílio na transferência e importação dos dados, suporte na validação dos dados, bem como garantia da qualidade dos dados, ou seja, caso seja verificada alguma inconsistência, é necessário que a contratada dê todo o suporte para que este problema seja resolvido, até mesmo via atualização dos dados, caso necessário. Com isso, a contratada deverá disponibilizar profissionais especializados, de segunda a sexta-feira, durante o horário comercial (das 08:00 às 12:00 e das 14:00 às 18:00), que tenham conhecimento técnico para esclarecimentos de dúvidas que venham a surgir durante o uso dos dados pela contratante. É esperado um tempo máximo para o esclarecimento de 48 horas após o envio da dúvida.