Datathon

Instruções para os participantes

O DESAFIO

O preço certo de um contrato?

Para disputar uma licitação, o Grupo HP precisa projetar quanto custará prestar o serviço durante toda a vigência do contrato. Esse preço depende de vários fatores, e os mais importantes são a quantidade de integrantes, sua rotatividade e sua escala de trabalho.

Vocês vão trabalhar com **dados dos integrantes do Grupo HP**. O desafio tem **três trilhas a serem concluídas**. Elas tratam de problemas diferentes, mas se conectam: cada uma estima uma parte do custo que compõe o preço do contrato. Considerem uma licitação com duração de três anos.

AS TRILHAS

Trilha 1: Rotatividade

O objetivo desta trilha é encontrar a melhor estratégia para projetar a rotatividade por empresa ao longo da vigência do contrato.

Considerem o turnover como fórmula central para medir a rotatividade:

$$\text{Turnover (\%)} = [\ (\text{Admissões} + \text{Rescisões}) / 2\ ] / \text{Headcount} \times 100$$

- **Admissões** é a quantidade de integrantes que entraram ao final do período.
- **Rescisões** é a quantidade de integrantes que saíram ao final do período.
- **Headcount** é a quantidade de integrantes com vínculo durante o período.

Trilha 2: Folguistas e feristas

O objetivo desta trilha é encontrar a melhor estratégia para metrificar a quantidade de integrantes a mais que cada empresa precisa para cobrir ausências por motivo de férias e atestados sem descumprir o contrato.

O serviço contratado precisa ser prestado todos os dias previstos, mesmo quando o integrante titular está ausente. Faltas, atestados, afastamentos e férias retiram integrantes da operação, e cada ausência precisa ser coberta por outro integrante. Nem toda ausência se comporta da mesma forma: algumas se planejam com antecedência, outras surgem sem aviso. Como transformar esse histórico em um quadro que mantenha o serviço sem sobrar nem faltar integrantes?

Trilha 3: Escala 6x1

O objetivo desta trilha é encontrar a melhor estratégia para metrificar a quantidade de integrantes a mais que cada empresa precisa se a escala mudar de 6x1 para 5x2.

Na escala 6x1 o integrante trabalha seis dias e folga um; na 5x2, trabalha cinco e folga dois. Com mais folgas por integrante, manter o mesmo serviço exige um quadro maior, e a mudança também pode alterar a rotatividade.

OS DADOS

O que os dados registram

Os dados contam a trajetória dos integrantes no Grupo HP em três partes:

- **Cadastro de vínculos.** Cada vínculo é um contrato de trabalho, do dia da admissão ao dia da demissão. O integrante que sai e é recontratado tem mais de um vínculo. O cadastro informa a função (a atividade exercida, como motorista de ônibus ou cobrador), o cargo padronizado (que agrupa funções equivalentes), a seção onde o integrante está lotado, o horário de trabalho, o salário e o motivo da saída, classificada como voluntária (quando o integrante pede para sair) ou involuntária (todos os outros casos, como decisão da empresa, fim de contrato ou falecimento).
- **Ponto diário.** É o registro da jornada de cada integrante em cada dia: quantos minutos deveria trabalhar segundo seu horário (zero nos dias de folga), quantos trabalhou de fato, quantos deixou de trabalhar por falta ou atraso, quanto dessa ausência foi abonado (justificado e aceito pela empresa, por exemplo com atestado médico) e quanto fez de hora extra. Cada dia também traz a situação do integrante naquela data: ativo, de férias, afastado etc.
- **Afastamentos.** São os períodos em que o integrante fica fora do trabalho sem perder o vínculo, como férias, licenças, atestados de doença e acidentes de trabalho, com data de início e de fim. Atestados de até 15 dias são pagos pela empresa; acima disso, pelo INSS.

O horário de cada integrante define sua escala de trabalho, como 6x1, 5x2 ou 12x36 (doze horas de trabalho e trinta e seis de descanso).

Período analisado

O ponto diário e os afastamentos vão de **01/01/2025 a 30/09/2026** (21 meses). O cadastro de vínculos traz todo o histórico, com admissões desde 1969 e demissões até 13/10/2026.

Tabelas

Tabela	O que é	Registros	Período
`dim_integrante`	Cadastro de vínculos: empresa, filial, seção, função, horário, sexo, escolaridade, faixa etária, salário de referência, situação, datas, tipo e motivo de admissão e categoria do desligamento	20.001 vínculos de 18.617 integrantes	Admissões de 1969 a 2026
`fBatidas`	Ponto diário: uma linha por vínculo por dia, com minutos previstos, de falta, de atraso, de abono, trabalhados e de hora extra, além de feriado e situação no dia	601.474 linhas	01/01/2025 a 30/09/2026
`fAfastamentos`	Férias, afastamentos (doença, INSS, acidente de trabalho, licenças) e desligamentos, com início e fim	5.052 eventos	Ativos entre 01/01/2025 e 30/09/2026
`dim_funcao`	Funções, com nome e cargo padronizado	10.175 funções	–
`dim_horario`	Horários de trabalho, com descrição, escala (5x2, 6x1, 12x36...), turno e tipo de jornada	1.368 horários	–
`dim_secao`	Seções, com empresa, filial, centro de custo, equipe e calendário	796 seções	–
`dim_empresa`	Filiais, com empresa e UF	13 filiais	–
`dim_feriados`	Feriados e pontos facultativos	32 feriados	2025 e 2026
`dim_situacao`	Descrição das situações (ativo, demitido, férias, afastado...)	11	–
`dim_grau_instrucao`	Descrição da escolaridade	12	–
`dim_tipo_admissao`, `dim_motivo_admissao`	Descrição do tipo e do motivo de admissão	6 e 3	–
`dim_tipo_funcionario`	Descrição do tipo de vínculo contratual	8	–

Acesso aos dados

Os dados estão disponíveis em: https://drive.google.com/drive/u/2/folders/1lCtH1PQDhBsrof7W0fUvcW8Lr9bG7nxW