

Teste do operador sample do MonetDB aplicado aos dados massivos do Bolsa Família

Roberto Mourão, Igor Stemler, Filipe Guedes

Orientadora: Prof^a Maristela Hollanda

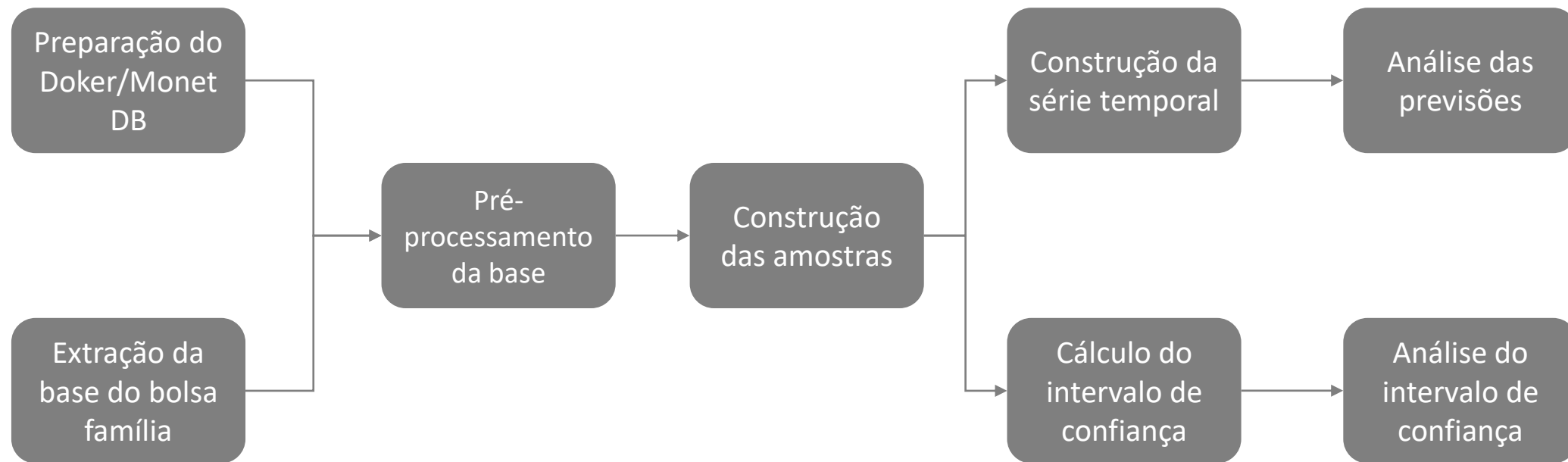
AGENDA

Trabalhos relacionados
Metodologia
Desenvolvimento do trabalho
Trabalhos
Conclusões

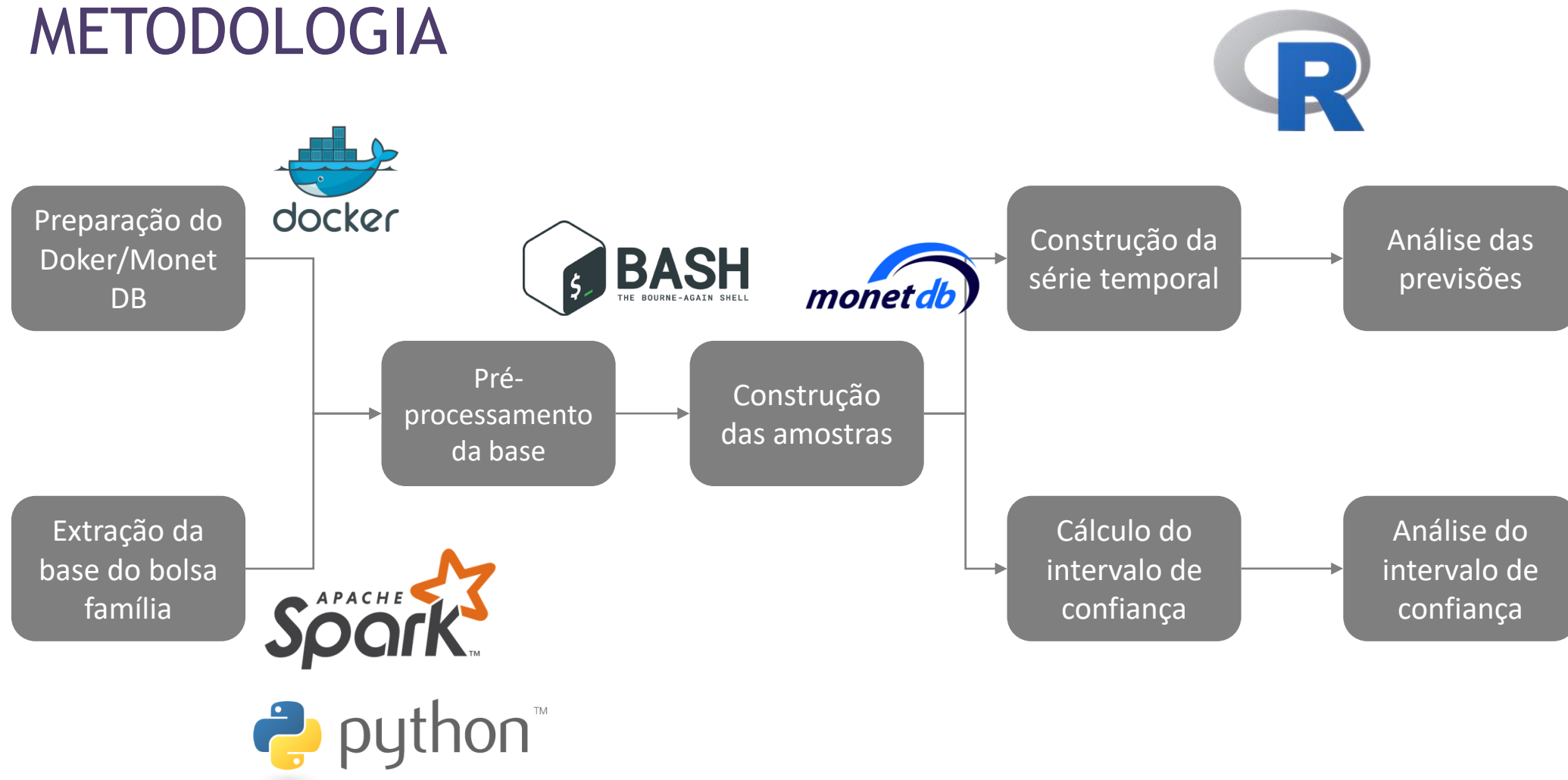
OBJETIVOS DO TRABALHO

TRABALHOS RELACIONADOS

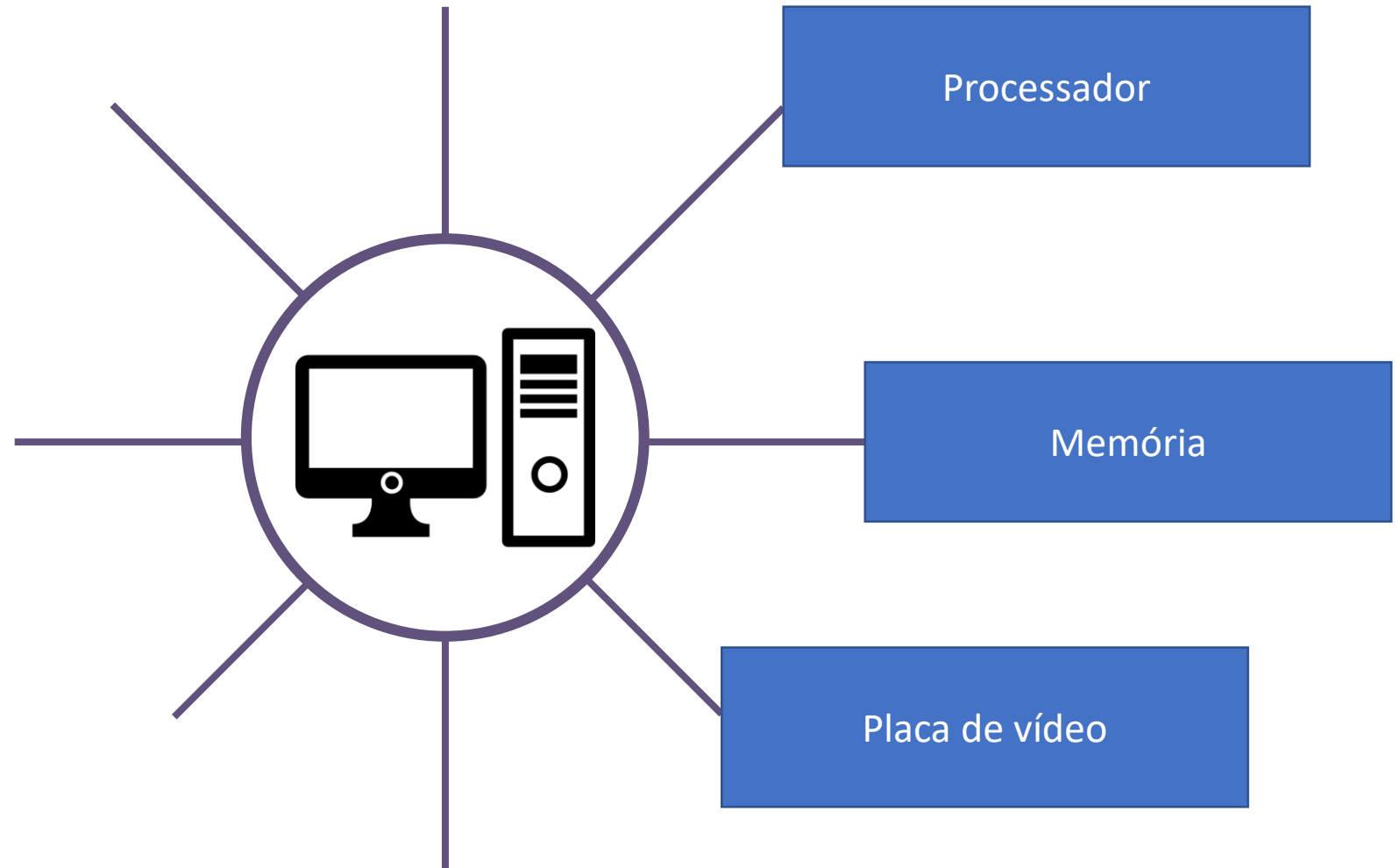
METODOLOGIA



METODOLOGIA



ESPECIFICAÇÃO DA MÁQUINA



DESENVOLVIMENTO DO TRABALHO

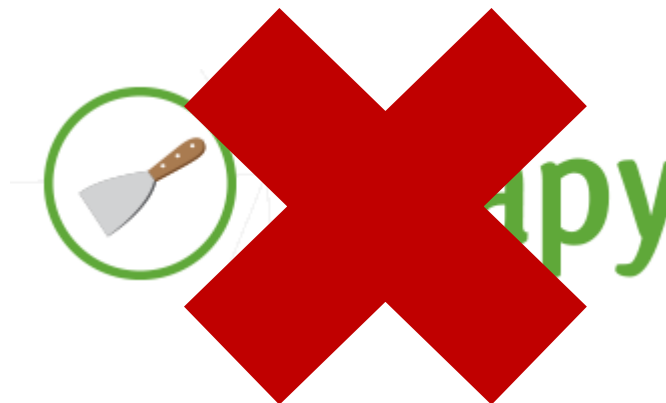
Extração da
base do bolsa
família

Ministério da Transparência, Fiscalização e Controladoria-Geral da União

Portal da Transparência

GOVERNO FEDERAL

Dados 2011 a 2017



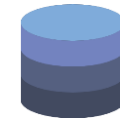
DESENVOLVIMENTO DO TRABALHO

Preparação do
Doker/Monet
DB

Cenário A – 1 Docker



Cenário B – 7 Dockers



DESENVOLVIMENTO DO TRABALHO

Pré-
processamento
da base

Etapa 1: Ajuste separador de milhar
Ajuste do campo data

- Etapa 2: Seleção das variáveis de interesse (data pagamento, cpf, valor, UF)
Criação de novos campos: entrada e saída

DESENVOLVIMENTO DO TRABALHO

Construção
das amostras

```
87  str
88  SAMPLEuniform_dbl(bat *r, bat *b, dbl *p) {
89      BAT *bb;
90      double pr = *p;
91      lng s;
92
93      if ( pr < 0.0 || pr > 1.0 ) {
94          throw(MAL, "sample.subuniform", ILLEGAL_ARGUMENT
95              " p should be between 0 and 1.0" );
96      } else if (pr == 0) { /* special case */
97          s = 0;
98          return SAMPLEuniform(r, b, &s);
99      }
100     if ((bb = BATdescriptor(*b)) == NULL) {
101         throw(MAL, "sample.subuniform", INTERNAL_BAT_ACCESS);
102     }
103     s = (lng) (pr*(double)BATcount(bb));
104     BBPunfix(bb->batCacheid);
105     return SAMPLEuniform(r, b, &s);
106 }
```

Falar da amostra de mil e 390 mil e justificar

DESENVOLVIMENTO DO TRABALHO

Construção da
série temporal

DESENVOLVIMENTO DO TRABALHO

Análise das
previsões

DESENVOLVIMENTO DO TRABALHO

Cálculo do
intervalo de
confiança

DESENVOLVIMENTO DO TRABALHO

Análise do
intervalo de
confiança

SUGESTÕES DE TRABALHOS FUTUROS

Comparar performance do monetDB com outros bancos de dados noSQL

CONCLUSÕES