

# Software Systems as Cities: A Controlled Experiment

MAC5779 - Engenharia de Software Experimental

Higor Amario de Souza<sup>1</sup>

<sup>1</sup>Instituto de Matemática e Estatística  
Universidade de São Paulo  
Programa de Pós-Graduação em Ciência da Computação

13 de Maio de 2013

# Introdução

- Visualização de software;
- Suporte a atividades de Engenharia de Software;
- Pouca validação empírica;
- Dados insuficientes para validação de hipótese.

# Introdução

- Visualização de software;
- Suporte a atividades de Engenharia de Software;
- Pouca validação empírica;
- Dados insuficientes para validação de hipótese.

# Introdução

- Visualização de software;
- Suporte a atividades de Engenharia de Software;
- Pouca validação empírica;
- Dados insuficientes para validação de hipótese.

# Introdução

- Visualização de software;
- Suporte a atividades de Engenharia de Software;
- Pouca validação empírica;
- Dados insuficientes para validação de hipótese.

## Outros trabalhos

Experiment	Subjects		Object system(s)	
	Academia	Industry	Classes	kLOC
Storey et al. [14]	30	0	17	2
Marcus et al. [6]	24	0	27	42
Lange et al. [3]	100	0	38 39	? ?
Quante [10]	25	0	475 1,725	43 160
Cornelissen et al. [1]	20	1	310	57
Wettel et al. [20]	21	20	1,320 4,656	93 454

# Lista de desejos de desenho experimental

- 1. Escolha um nível de comparação justo;
- 2. Inclua participantes da indústria;
- 3. Considere o nível de experiência dos participantes;
- 4. Forneça um tutorial da ferramenta de experimentação para os participantes;
- 5. Identifique um conjunto de tarefas relevantes;
- 6. Inclua tarefas que podem não ter vantagem com a ferramenta avaliada;

# Lista de desejos de desenho experimental

- 1. Escolha um nível de comparação justo;
- 2. Inclua participantes da indústria;
- 3. Considere o nível de experiência dos participantes;
- 4. Forneça um tutorial da ferramenta de experimentação para os participantes;
- 5. Identifique um conjunto de tarefas relevantes;
- 6. Inclua tarefas que podem não ter vantagem com a ferramenta avaliada;

# Lista de desejos de desenho experimental

- 1. Escolha um nível de comparação justo;
- 2. Inclua participantes da indústria;
- 3. Considere o nível de experiência dos participantes;
- 4. Forneça um tutorial da ferramenta de experimentação para os participantes;
- 5. Identifique um conjunto de tarefas relevantes;
- 6. Inclua tarefas que podem não ter vantagem com a ferramenta avaliada;

## Lista de desejos de desenho experimental

- 1. Escolha um nível de comparação justo;
- 2. Inclua participantes da indústria;
- 3. Considere o nível de experiência dos participantes;
- 4. Forneça um tutorial da ferramenta de experimentação para os participantes;
- 5. Identifique um conjunto de tarefas relevantes;
- 6. Inclua tarefas que podem não ter vantagem com a ferramenta avaliada;

## Lista de desejos de desenho experimental

- 1. Escolha um nível de comparação justo;
- 2. Inclua participantes da indústria;
- 3. Considere o nível de experiência dos participantes;
- 4. Forneça um tutorial da ferramenta de experimentação para os participantes;
- 5. Identifique um conjunto de tarefas relevantes;
- 6. Inclua tarefas que podem não ter vantagem com a ferramenta avaliada;

# Lista de desejos de desenho experimental

- 1. Escolha um nível de comparação justo;
- 2. Inclua participantes da indústria;
- 3. Considere o nível de experiência dos participantes;
- 4. Forneça um tutorial da ferramenta de experimentação para os participantes;
- 5. Identifique um conjunto de tarefas relevantes;
- 6. Inclua tarefas que podem não ter vantagem com a ferramenta avaliada;

# Lista de desejos de desenho experimental

- 7. Limite o tempo para resolver cada tarefa;
- 8. Escolha sistemas do mundo real;
- 9. Inclua mais de um sistema no experimento;
- 10. Forneça os mesmos dados para todos os participantes;
- 11. Relate os resultados nas tarefas individuais;
- 12. Forneça os detalhes necessários para tornar o experimento replicável.

# Lista de desejos de desenho experimental

- 7. Limite o tempo para resolver cada tarefa;
- 8. Escolha sistemas do mundo real;
- 9. Inclua mais de um sistema no experimento;
- 10. Forneça os mesmos dados para todos os participantes;
- 11. Relate os resultados nas tarefas individuais;
- 12. Forneça os detalhes necessários para tornar o experimento replicável.

# Lista de desejos de desenho experimental

- 7. Limite o tempo para resolver cada tarefa;
- 8. Escolha sistemas do mundo real;
- 9. Inclua mais de um sistema no experimento;
- 10. Forneça os mesmos dados para todos os participantes;
- 11. Relate os resultados nas tarefas individuais;
- 12. Forneça os detalhes necessários para tornar o experimento replicável.

## Lista de desejos de desenho experimental

- 7. Limite o tempo para resolver cada tarefa;
- 8. Escolha sistemas do mundo real;
- 9. Inclua mais de um sistema no experimento;
- 10. Forneça os mesmos dados para todos os participantes;
- 11. Relate os resultados nas tarefas individuais;
- 12. Forneça os detalhes necessários para tornar o experimento replicável.

## Lista de desejos de desenho experimental

- 7. Limite o tempo para resolver cada tarefa;
- 8. Escolha sistemas do mundo real;
- 9. Inclua mais de um sistema no experimento;
- 10. Forneça os mesmos dados para todos os participantes;
- 11. Relate os resultados nas tarefas individuais;
- 12. Forneça os detalhes necessários para tornar o experimento replicável.

## Lista de desejos de desenho experimental

- 7. Limite o tempo para resolver cada tarefa;
- 8. Escolha sistemas do mundo real;
- 9. Inclua mais de um sistema no experimento;
- 10. Forneça os mesmos dados para todos os participantes;
- 11. Relate os resultados nas tarefas individuais;
- 12. Forneça os detalhes necessários para tornar o experimento replicável.

# Desenho experimental

- Comparação entre a ferramenta e o estado da prática;
- Questões:
  - Medir correção e tempo;
  - Tipos de tarefas com maior adequação;
  - Diferença entre praticantes da indústria e academia;
  - Diferença entre praticantes novatos e experientes.
- Variáveis independentes:
  - Ferramenta e tamanho do programa;
- Variáveis dependentes:
  - Medir correção e tempo;
- Dois grupos apenas pós-teste.

# Desenho experimental

- Comparação entre a ferramenta e o estado da prática;
- Questões:
  - Medir correção e tempo;
  - Tipos de tarefas com maior adequação;
  - Diferença entre praticantes da indústria e academia;
  - Diferença entre praticantes novatos e experientes.
- Variáveis independentes:
  - Ferramenta e tamanho do programa;
- Variáveis dependentes:
  - Medir correção e tempo;
- Dois grupos apenas pós-teste.

# Desenho experimental

- Comparação entre a ferramenta e o estado da prática;
- Questões:
  - Medir correção e tempo;
  - Tipos de tarefas com maior adequação;
  - Diferença entre praticantes da indústria e academia;
  - Diferença entre praticantes novatos e experientes.
- Variáveis independentes:
  - Ferramenta e tamanho do programa;
- Variáveis dependentes:
  - Medir correção e tempo;
- Dois grupos apenas pós-teste.

# Desenho experimental

- Comparação entre a ferramenta e o estado da prática;
- Questões:
  - Medir correção e tempo;
  - Tipos de tarefas com maior adequação;
  - Diferença entre praticantes da indústria e academia;
  - Diferença entre praticantes novatos e experientes.
- Variáveis independentes:
  - Ferramenta e tamanho do programa;
- Variáveis dependentes:
  - Medir correção e tempo;
- Dois grupos apenas pós-teste.

# Desenho experimental

- Comparação entre a ferramenta e o estado da prática;
- Questões:
  - Medir correção e tempo;
  - Tipos de tarefas com maior adequação;
  - Diferença entre praticantes da indústria e academia;
  - Diferença entre praticantes novatos e experientes.
- Variáveis independentes:
  - Ferramenta e tamanho do programa;
- Variáveis dependentes:
  - Medir correção e tempo;
- Dois grupos apenas pós-teste.

# Desenho experimental

- Comparação entre a ferramenta e o estado da prática;
- Questões:
  - Medir correção e tempo;
  - Tipos de tarefas com maior adequação;
  - Diferença entre praticantes da indústria e academia;
  - Diferença entre praticantes novatos e experientes.
- Variáveis independentes:
  - Ferramenta e tamanho do programa;
- Variáveis dependentes:
  - Medir correção e tempo;
- Dois grupos apenas pós-teste.

# Desenho experimental

- Comparação entre a ferramenta e o estado da prática;
- Questões:
  - Medir correção e tempo;
  - Tipos de tarefas com maior adequação;
  - Diferença entre praticantes da indústria e academia;
  - Diferença entre praticantes novatos e experientes.
- Variáveis independentes:
  - Ferramenta e tamanho do programa;
- Variáveis dependentes:
  - Medir correção e tempo;
- Dois grupos apenas pós-teste.

# Desenho experimental

- Comparação entre a ferramenta e o estado da prática;
- Questões:
  - Medir correção e tempo;
  - Tipos de tarefas com maior adequação;
  - Diferença entre praticantes da indústria e academia;
  - Diferença entre praticantes novatos e experientes.
- Variáveis independentes:
  - Ferramenta e tamanho do programa;
- Variáveis dependentes:
  - Medir correção e tempo;
- Dois grupos apenas pós-teste.

# Desenho experimental

- Comparação entre a ferramenta e o estado da prática;
- Questões:
  - Medir correção e tempo;
  - Tipos de tarefas com maior adequação;
  - Diferença entre praticantes da indústria e academia;
  - Diferença entre praticantes novatos e experientes.
- Variáveis independentes:
  - Ferramenta e tamanho do programa;
- Variáveis dependentes:
  - Medir correção e tempo;
  - Dois grupos apenas pós-teste.

# Desenho experimental

- Comparação entre a ferramenta e o estado da prática;
- Questões:
  - Medir correção e tempo;
  - Tipos de tarefas com maior adequação;
  - Diferença entre praticantes da indústria e academia;
  - Diferença entre praticantes novatos e experientes.
- Variáveis independentes:
  - Ferramenta e tamanho do programa;
- Variáveis dependentes:
  - Medir correção e tempo;
- Dois grupos apenas pós-teste.

# Desenho experimental

- Comparação entre a ferramenta e o estado da prática;
- Questões:
  - Medir correção e tempo;
  - Tipos de tarefas com maior adequação;
  - Diferença entre praticantes da indústria e academia;
  - Diferença entre praticantes novatos e experientes.
- Variáveis independentes:
  - Ferramenta e tamanho do programa;
- Variáveis dependentes:
  - Medir correção e tempo;
- Dois grupos apenas pós-teste.

# Desenho experimental

Ferramentas do estado da prática:

- Compreensão do programa: Eclipse;
- Avaliação da qualidade do design: Excel;
- Análise de evolução: Excluído;

Programas:

- Findbugs;
- Azureus.

# Desenho experimental

Ferramentas do estado da prática:

- Compreensão do programa: Eclipse;
- Avaliação da qualidade do design: Excel;
- Análise de evolução: Excluído;

Programas:

- Findbugs;
- Azureus.

# Desenho experimental

Ferramentas do estado da prática:

- Compreensão do programa: Eclipse;
- Avaliação da qualidade do design: Excel;
- Análise de evolução: Excluído;

Programas:

- Findbugs;
- Azureus.

# Desenho experimental

Ferramentas do estado da prática:

- Compreensão do programa: Eclipse;
- Avaliação da qualidade do design: Excel;
- Análise de evolução: Excluído;

Programas:

- Findbugs;
- Azureus.

# Desenho experimental

Ferramentas do estado da prática:

- Compreensão do programa: Eclipse;
- Avaliação da qualidade do design: Excel;
- Análise de evolução: Excluído;

Programas:

- Findbugs;
- Azureus.

# Operação

- Piloto;
- Treinamento.

# Operação

- Piloto;
- Treinamento.

# Coleta de dados

- 10 minutos por tarefa;
- Feedback.

# Coleta de dados

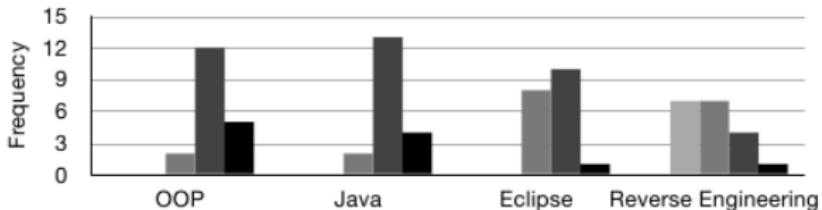
- 10 minutos por tarefa;
- Feedback.

# Análise

Block			Treatment				Total
			T1	T2	T3	T4	
Block	Academia	Beginner	2	3	2	2	9
		Advanced	2	2	3	5	12
	Industry	Beginner	0	0	0	0	0
		Advanced	6	7	3	4	20
Total			10	12	8	11	41

# Análise

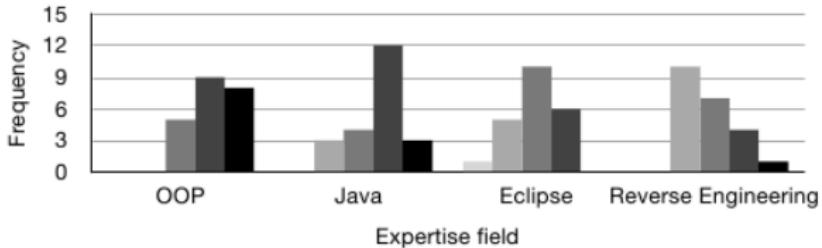
Ecl&Exl



Legend:

- none
- beginner
- knowledgeable
- advanced
- expert

CodeCity



Legend:

- none
- beginner
- knowledgeable
- advanced
- expert

# Resultados

- Análise de Variância (ANOVA);
- Distribuição normal;

Dep. var. System size Tool	Correctness (points)						Completion Time (minutes)					
	Medium		Large		Overall		Medium		Large		Overall	
	Ecl+Exl	CodeCity	Ecl+Exl	CodeCity	Ecl+Exl	CodeCity	Ecl+Exl	CodeCity	Ecl+Exl	CodeCity	Ecl+Exl	CodeCity
mean	5.462	6.733	3.896	5.050	4.803	5.968	38.809	33.178	44.128	39.644	41.048	36.117
difference		+23.27%		+29.62%		+24.26%		-14.51%		-10.16%		-12.01%
min	3.50	5.00	2.27	3.00	2.27	3.00	31.92	24.67	22.83	27.08	22.83	24.67
max	6.50	8.00	6.00	6.30	6.50	8.00	53.08	39.50	55.92	48.55	55.92	48.55
median	5.800	6.585	3.900	5.100	4.430	6.065	38.000	35.575	48.260	40.610	40.080	36.125
stdev	1.147	.959	1.085	1.031	1.349	1.294	6.789	5.545	11.483	6.963	9.174	6.910

# Resultados

- Análise de Variância (ANOVA);
- Distribuição normal;

Dep. var. System size Tool	Correctness (points)						Completion Time (minutes)					
	Medium		Large		Overall		Medium		Large		Overall	
	Ecl+Exl	CodeCity	Ecl+Exl	CodeCity	Ecl+Exl	CodeCity	Ecl+Exl	CodeCity	Ecl+Exl	CodeCity	Ecl+Exl	CodeCity
mean	5.462	6.733	3.896	5.050	4.803	5.968	38.809	33.178	44.128	39.644	41.048	36.117
difference		+23.27%		+29.62%		+24.26%		-14.51%		-10.16%		-12.01%
min	3.50	5.00	2.27	3.00	2.27	3.00	31.92	24.67	22.83	27.08	22.83	24.67
max	6.50	8.00	6.00	6.30	6.50	8.00	53.08	39.50	55.92	48.55	55.92	48.55
median	5.800	6.585	3.900	5.100	4.430	6.065	38.000	35.575	48.260	40.610	40.080	36.125
stdev	1.147	.959	1.085	1.031	1.349	1.294	6.789	5.545	11.483	6.963	9.174	6.910

# Resultados

- Análise de Variância (ANOVA);
- Distribuição normal;

Dep. var. System size Tool	Correctness (points)						Completion Time (minutes)					
	Medium		Large		Overall		Medium		Large		Overall	
	Ecl+Exl	CodeCity	Ecl+Exl	CodeCity	Ecl+Exl	CodeCity	Ecl+Exl	CodeCity	Ecl+Exl	CodeCity	Ecl+Exl	CodeCity
mean	5.462	6.733	3.896	5.050	4.803	5.968	38.809	33.178	44.128	39.644	41.048	36.117
difference		+23.27%		+29.62%		+24.26%		-14.51%		-10.16%		-12.01%
min	3.50	5.00	2.27	3.00	2.27	3.00	31.92	24.67	22.83	27.08	22.83	24.67
max	6.50	8.00	6.00	6.30	6.50	8.00	53.08	39.50	55.92	48.55	55.92	48.55
median	5.800	6.585	3.900	5.100	4.430	6.065	38.000	35.575	48.260	40.610	40.080	36.125
stdev	1.147	.959	1.085	1.031	1.349	1.294	6.789	5.545	11.483	6.963	9.174	6.910

# Resultados

## Correção:

- Não existe correlação entre correção e tamanho do programa;
- CodeCity significativamente maior;
- Tamanho do programa: médio foi significativamente maior;
- 24.26%;

## Tempo:

- Não existe correlação entre tempo e tamanho do programa;
- CodeCity significativamente menor;
- Tamanho do programa: médio foi significativamente menor;
- 12.01%.

# Resultados

## Correção:

- Não existe correlação entre correção e tamanho do programa;
- CodeCity significativamente maior;
- Tamanho do programa: médio foi significativamente maior;
- 24.26%;

## Tempo:

- Não existe correlação entre tempo e tamanho do programa;
- CodeCity significativamente menor;
- Tamanho do programa: médio foi significativamente menor;
- 12.01%.

# Resultados

## Correção:

- Não existe correlação entre correção e tamanho do programa;
- CodeCity significativamente maior;
- Tamanho do programa: médio foi significativamente maior;
- 24.26%;

## Tempo:

- Não existe correlação entre tempo e tamanho do programa;
- CodeCity significativamente menor;
- Tamanho do programa: médio foi significativamente menor;
- 12.01%.

# Resultados

## Correção:

- Não existe correlação entre correção e tamanho do programa;
- CodeCity significativamente maior;
- Tamanho do programa: médio foi significativamente maior;
- 24.26%;

## Tempo:

- Não existe correlação entre tempo e tamanho do programa;
- CodeCity significativamente menor;
- Tamanho do programa: médio foi significativamente menor;
- 12.01%.

# Resultados

## Correção:

- Não existe correlação entre correção e tamanho do programa;
- CodeCity significativamente maior;
- Tamanho do programa: médio foi significativamente maior;
- 24.26%;

## Tempo:

- Não existe correlação entre tempo e tamanho do programa;
- CodeCity significativamente menor;
- Tamanho do programa: médio foi significativamente menor;
- 12.01%.

# Resultados

## Correção:

- Não existe correlação entre correção e tamanho do programa;
- CodeCity significativamente maior;
- Tamanho do programa: médio foi significativamente maior;
- 24.26%;

## Tempo:

- Não existe correlação entre tempo e tamanho do programa;
- CodeCity significativamente menor;
- Tamanho do programa: médio foi significativamente menor;
- 12.01%.

# Resultados

## Correção:

- Não existe correlação entre correção e tamanho do programa;
- CodeCity significativamente maior;
- Tamanho do programa: médio foi significativamente maior;
- 24.26%;

## Tempo:

- Não existe correlação entre tempo e tamanho do programa;
- CodeCity significativamente menor;
- Tamanho do programa: médio foi significativamente menor;
- 12.01%.

## Análise das tarefas

- Vantagens no tempo pela visão geral de forma visual;
- Uma lista com o maior valor é verificada mais rapidamente que usando a CodeCity;
- Combinar as estratégias (lista + CodeCity).

## Análise das tarefas

- Vantagens no tempo pela visão geral de forma visual;
- Uma lista com o maior valor é verificada mais rapidamente que usando a CodeCity;
- Combinar as estratégias (lista + CodeCity).

## Análise das tarefas

- Vantagens no tempo pela visão geral de forma visual;
- Uma lista com o maior valor é verificada mais rapidamente que usando a CodeCity;
- Combinar as estratégias (lista + CodeCity).

# Ameaças à validade

## Internas:

- Tarefas que favorecem CodeCity;
- Tempo por tarefa;
- Comparação entre modelo e código-fonte;
- Treinamento somente para CodeCity;

## Externas:

- Tarefas podem não refletir situações reais de engenharia reversa;
- Programas podem não ser adequados para as tarefas;
- Viés do experimentador.

# Ameaças à validade

## Internas:

- Tarefas que favorecem CodeCity;
- Tempo por tarefa;
- Comparação entre modelo e código-fonte;
- Treinamento somente para CodeCity;

## Externas:

- Tarefas podem não refletir situações reais de engenharia reversa;
- Programas podem não ser adequados para as tarefas;
- Viés do experimentador.

# Ameaças à validade

## Internas:

- Tarefas que favorecem CodeCity;
- Tempo por tarefa;
- Comparação entre modelo e código-fonte;
- Treinamento somente para CodeCity;

## Externas:

- Tarefas podem não refletir situações reais de engenharia reversa;
- Programas podem não ser adequados para as tarefas;
- Viés do experimentador.

# Ameaças à validade

## Internas:

- Tarefas que favorecem CodeCity;
- Tempo por tarefa;
- Comparação entre modelo e código-fonte;
- Treinamento somente para CodeCity;

## Externas:

- Tarefas podem não refletir situações reais de engenharia reversa;
- Programas podem não ser adequados para as tarefas;
- Viés do experimentador.

# Ameaças à validade

## Internas:

- Tarefas que favorecem CodeCity;
- Tempo por tarefa;
- Comparação entre modelo e código-fonte;
- Treinamento somente para CodeCity;

## Externas:

- Tarefas podem não refletir situações reais de engenharia reversa;
- Programas podem não ser adequados para as tarefas;
- Viés do experimentador.

# Ameaças à validade

## Internas:

- Tarefas que favorecem CodeCity;
- Tempo por tarefa;
- Comparação entre modelo e código-fonte;
- Treinamento somente para CodeCity;

## Externas:

- Tarefas podem não refletir situações reais de engenharia reversa;
- Programas podem não ser adequados para as tarefas;
- Viés do experimentador.

# Observações

- Como foram escolhidos os participantes?
- Porque o número de participantes foi considerado suficiente?
- A diferença entre os programas Findbugs e Azureus pode estar relacionada com a dificuldade de compreensão dos programas (ameaça de instrumentação);
- Como medir a vantagem da abordagem visual?

# Observações

- Como foram escolhidos os participantes?
- Porque o número de participantes foi considerado suficiente?
- A diferença entre os programas Findbugs e Azureus pode estar relacionada com a dificuldade de compreensão dos programas (ameaça de instrumentação);
- Como medir a vantagem da abordagem visual?

# Observações

- Como foram escolhidos os participantes?
- Porque o número de participantes foi considerado suficiente?
- A diferença entre os programas Findbugs e Azureus pode estar relacionada com a dificuldade de compreensão dos programas (ameaça de instrumentação);
- Como medir a vantagem da abordagem visual?

# Observações

- Como foram escolhidos os participantes?
- Porque o número de participantes foi considerado suficiente?
- A diferença entre os programas Findbugs e Azureus pode estar relacionada com a dificuldade de compreensão dos programas (ameaça de instrumentação);
- Como medir a vantagem da abordagem visual?

# Referência

- Software Systems as Cities: A Controlled Experiment  
Richard Wettel, Michael Lanza & Romain Robbes  
Proceedings of the 33rd International Conference on Software Engineering P.551-560  
2011.