

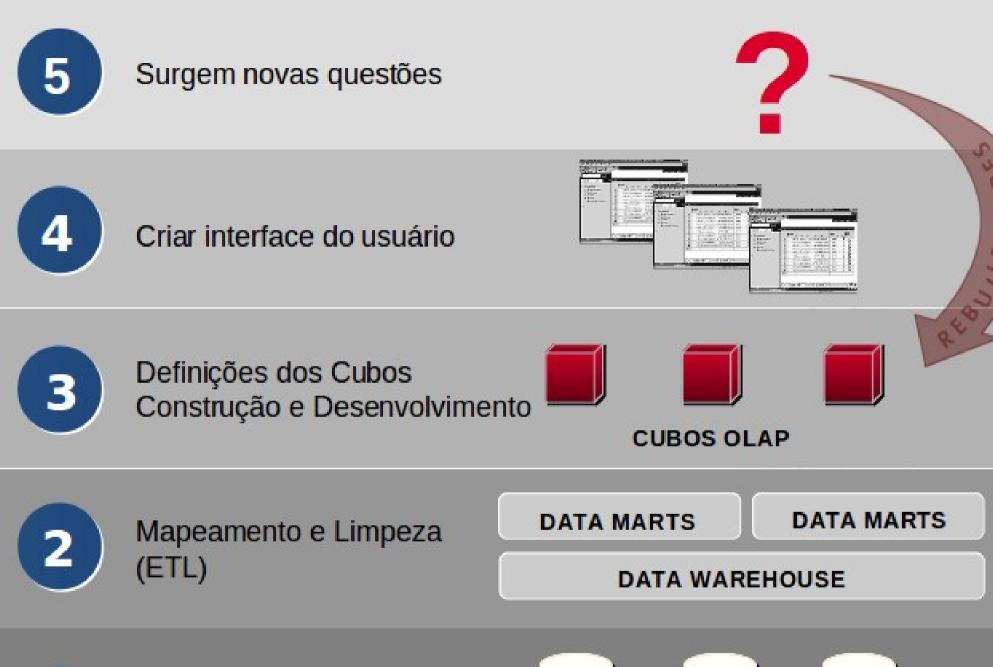
Rafaela Raganham

- •→ Especialista em Business Intelligence, desenvolvedora de software. Atua pela Ambiente Livre com analises, implantação, instrução e customização de Software Livres tais comoPentaho, Hadoop, LimeSurvey, SugarCRM e GLPI.
 - → Formada em Sistemas de Informação pela Unibrasil.
 - → MBA Business Intelligence pela Universidade Positivo.
- → Instrutora nos treinamentos de Pentaho, Limesurvey e SugarCRM.
- → Professora Universitária da Universidade Positivo no Curso de MBA em Business Intelligence
- → Palestrante em Eventos nacionais e Internacionais tais como FISL, PentahoDay, Latinoware, FTSL, SFD e SOLISC.

Business Intelligence

- Conjunto de conceitos, métodos e recursos tecnológicos que habilitam a obtenção e distribuição de informações geradas a partir de dados. Operacionais e históricos visando gerar resultados para tomada de decisão;
- O termo Business Intelligence (BI), pode ser traduzido como Inteligência de negócios, refere-se ao processo de coleta, organização, análise, compartilhamento e monitoramento de informações que oferecem suporte a gestão de negócios.





Selecionar bases de dados



 A Pentaho é uma empresa de origem norte-americana, sediada na Florida, EUA desenvolvedora do Pentaho.

- Integrou projetos open source maduros e contratou alguns dos principais desenvolvedores.
- Desenvolve novas funcionalidades para versão CE e Enterprise.















Vantagens Pentaho Bl

- Custo zero de licença de software (CE).
- Portal WEB.
- Consolidação de Informações da empresa.
- Extrair e integrar dados de múltiplas fontes.
- Faz uso da experiência.
- Procurar relações de causa e efeito.
- Transformar os registros obtidos em informação útil para o conhecimento empresarial.

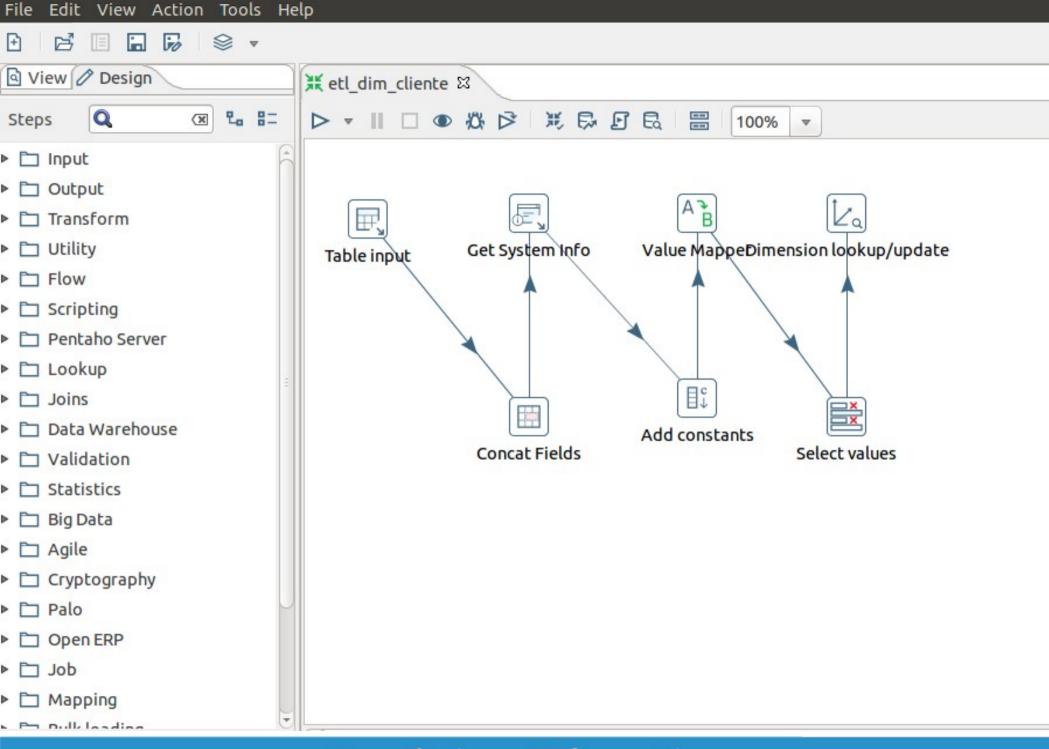


- → + de 8 Milhões de Downloads (média de 1 a cada 30 segundos)
- → Em mais de 185 países

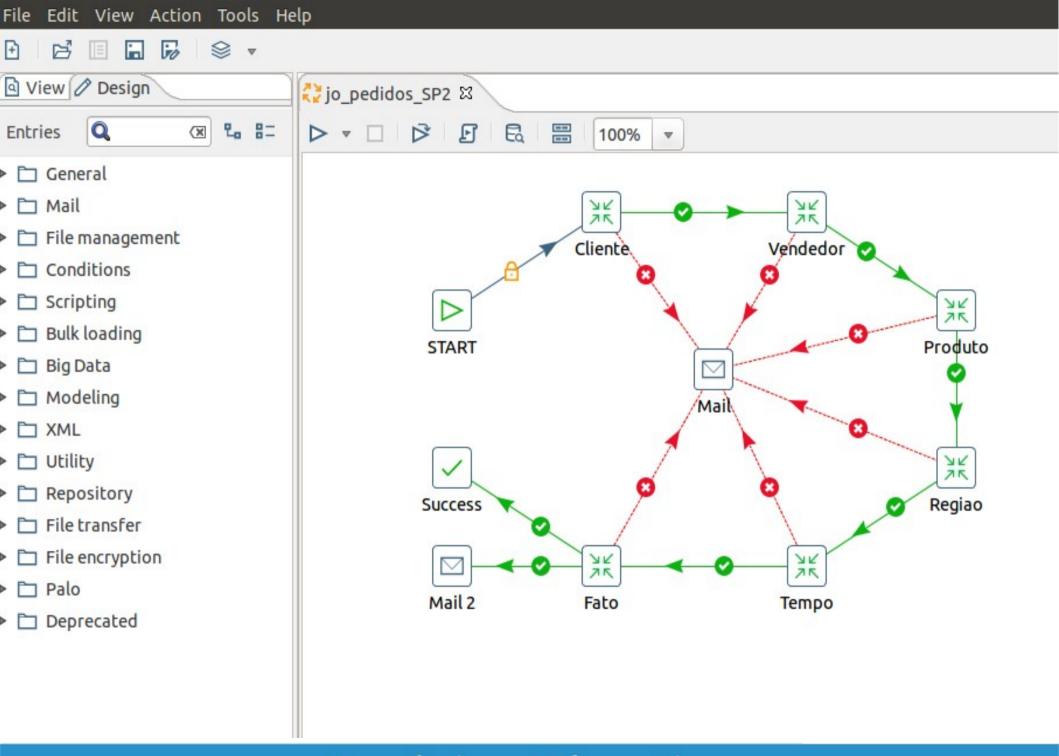


Pentaho Data Integration

- Rica biblioteca de transformação com mais de 180 objetos de mapeamento.
- Ampla fonte de dados, incluindo pacotes de aplicações, mais de 30 plataformas de código aberto e proprietário do banco de dados, planilha eletrônicas, etc..
- Suporte avançado de armazenagem de dados para Dimensões.



Tecnologia em Software Livre



Tecnologia em Software Livre

Pentaho Analysis

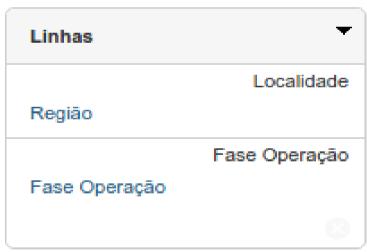
- Explorar livremente informações comerciais através de perfuração e tabulação de dados.
- Velocidade de resposta as consultas analíticascomplexas.
- Visualiza a informação multidimensional, escolhendo de métricas e atributos específicos para analisar

Pentaho Analysis

- Totalmente Web, drag-and-drop e criação de relatórios.
- Classificação e filtragem avançadas
- Totais personalizados e definidos pelo usuário cálculos
- Gráfico visualizações







Filtros		•

		ACIDENTE	INCIDENTE GRAVE
Região	Fase Operação	Quantidade	Quantidade
Centro-Oeste	DESCIDA	36	13
	SUBIDA	60	8
Nordeste	DESCIDA	102	14
	SUBIDA	36	10
Norte	DESCIDA	32	11
	SUBIDA	101	-
Sudeste	DESCIDA	87	14
	SUBIDA	184	34
Sul	DESCIDA	16	12
	SUBIDA	73	-
	Grand Total	-	-

Tecnologia em Software Livre

Quantidade Centro-Oeste / Sul / 18.1% . Nordeste / 18.9% Sudeste / 28.5% Quantidade 2,000 ¬ 1,900 -1,800 -1,700 -1,600 -1,500 -1,400 -1,300 -1,200 -1,100 -1,000 -900 -800 -



Centro-Oeste

Nordeste

Norte

Sudeste

Sul

700 -600 -500 -400 -300 -200 -

Pentaho Report Designer

- → Consulta totalmente Web
- → Amplo suporte de fonte de dados, incluindo relacionais, OLAP, XML.
- → Opções de saída Popularer (PDF, HTML, Microsoft Excel, Rich Text Format, ou texto simples)
- → Implantação flexível de relatórios incorporados ao BI.

Pentaho Report Designer

- → Design de relatórios gerador rapidamente com o assistente de relatório
- → Conectar-se a diversas fontes de dados, incluindo dados relacionais, Pentaho Analysis, arquivos flat, objetos Java, ou mesmo fluxo de dados transformações Pentaho Data Integration
- → Publique diretamente para o servidor de BI



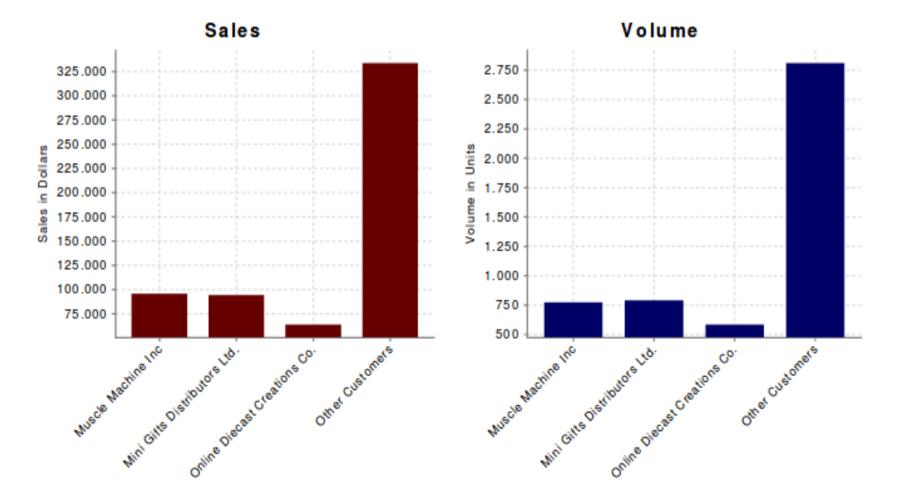
Classic Cars, Planes, Ships, Trains

			Classic Cars,	Pla	nes, a	onips,	ıraı					
INE: Classic Cars												
Vendor	BAR CODE	SKU	Name	Scale	On Hand	Cost	MSRP					
utoart Studio Design		S12_1099	1968 Ford Mustang	1:12	68 units	\$95	\$ 195					
	Description: Hood, doors and trunk all open to reveal highly detailed interior features. Steering wheel actually turns the front wheels. Color dark green.											
68												
arousel DieCast Legends		S24_1628	1986 Shelby Cobra 427 S/C	1:24	8.197 unit	\$29	\$ 50					
Description: This die cast model of the 1966 Shelby Cobra 427 S/C includes many authentic details and operating parts. The 1:24 scale model of this iconic lighweig sports can from the 1960s comes in silver and it's own display case.												
8.197												
		S24_2840	1958 Chevy Corvette Limited Edition	1:24	2.542 unit	\$16	\$ 35					
	Description: The operating parts of this 1958 Chevy Corvette Limited Edition are particularly delicate due to their precise scale and require special care and attention Features rotating wheels, working streering, opening doors and trunk. Color dark green.											
	2.542											
		S700_2824	1982 Camaro Z28	1:18	6.934 unit	\$47	\$ 101					
	Description: Features include openin Measures approximately 9 1/2 Long	ng and closing doors	. Color: White.									
			6.934									
lassic Metal Creations		S10_1949	1952 Alpine Renault 1300	1:10	7.305 unit	\$99	\$ 214					
	Description: Turnable front wheels;	steering function; det	alled interior; detailed engine; opening hood; opening	trunk; openi	ing doors; and	detailed chase	sis.					
			7.305									



Top 3 Customers

Customer	Sales (US \$)	Volume (Units)	% of Total
Muscle Machine Inc	\$ 95.657	774	16,28%
Mini Gifts Distributors Ltd.	\$ 94.211	791	16,04%
Online Diecast Creations Co.	\$ 63.981	585	10,89%
Other Customers	\$ 333.579	2.809	56,79%
Total	\$ 587.428	4.959	,



Pentaho Dashboard

- Oferecer métricas em uma atrativa interface intuitiva e visual, painéis proporcionam aos gestores empresariais a informação crítica de que necessitam para entender e melhorar o desempenho organizacional.
- Os Dashboards fornecem uma representação ilustrada do desempenho dos negócios em toda a organização

Theme Park Dashboard

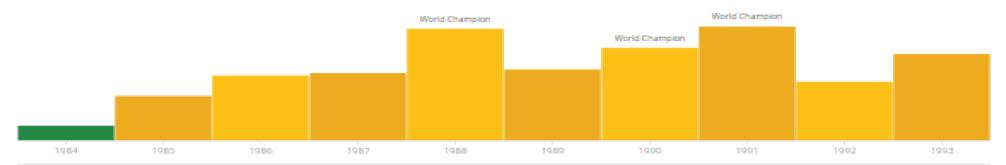
Welcome pentaho.



Total GPs	Race Wins	Podlums	Poles	Fastest Laps	Championships
162	41	81	65	19	3

Season Review

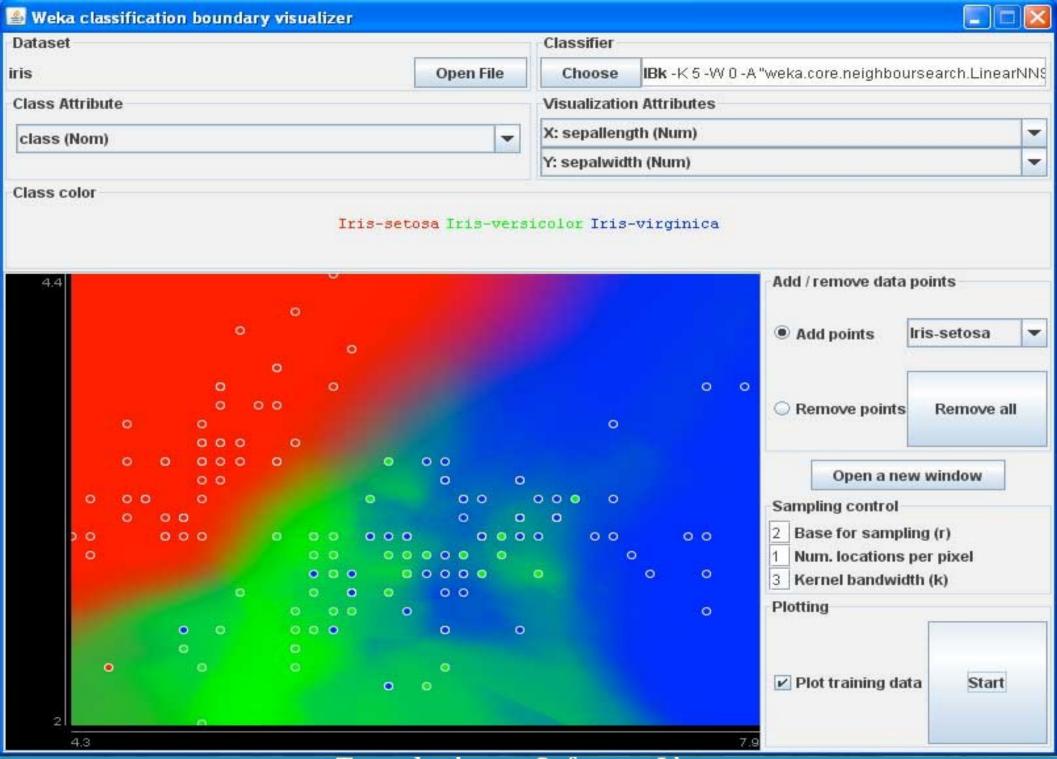
GP Points (click bars to select the season under analysis)



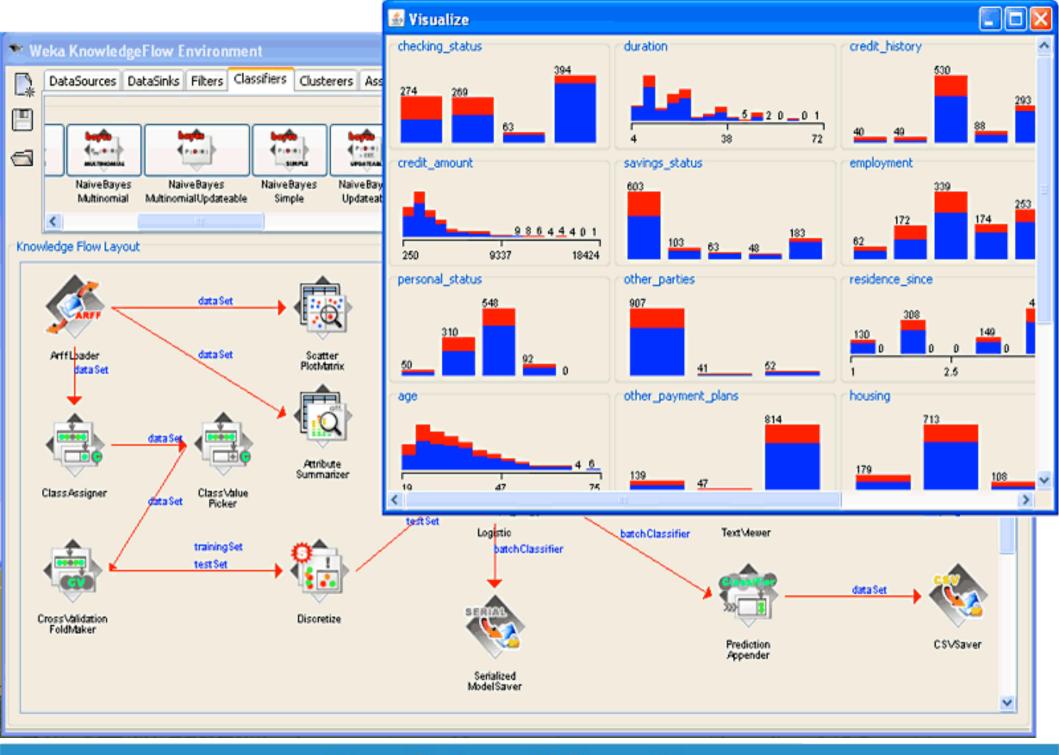
Circuit	20th	19th	18th	17%h	16th	15th	14th	13th	1.25h	11th	10th	9th	Bth	7th	6th	5th	4th	3rd	2nd	1st
Jacarepaguá					•															
Ryalami								•							5000					
Zolder		•													200					
Imola																				
Dijan								•												
Monaco								•											500	
G. Villeneuve												•		200						
Detroit														•						
Fair Park															•					
Brands Hatch														•				5000		
Hockenheim												•								
Spielberg											•									
Zandvoort								•												
Monza																				
Nürburgring									•											
Estorii																		* C 3*		

Data Mining - Weka

- → É um nível de excelência de Bl.
- → Fornece informações sobre padrões e relacionamentos ocultos nos seus dados.
- → Permite-lhe explorar correlações para melhorar o desempenho organizacional
- → Fornece indicadores de desempenho futuro.
- → Permite a incorporação de recomendações em suas aplicações
- → Permite-lhe tirar o máximo partido de um conjunto de algoritmos de mineração de dados



Tecnologia em Software Livre



Tecnologia em Software Livre



Dados Abertos



 Dados abertos são dados que podem ser livremente usados, reutilizados e redistribuídos por qualquer pessoa - sujeitos, no máximo, à exigência de atribuição da fonte e compartilhamento pelas mesmas regras.

→ Completo

Todos os dados públicos devem ser disponibilizados.

→ Primár<u>ios</u>

Os dados devem ser coletados na fonte com o maior nível de detalhamento possível

→ Oportunidade

Sua disponibilidade deve ser feita tão rapidamente quanto necessário para preservar o valor dos dados.

→ Acessibilidade

Os dados devem estar disponíveis para a mais ampla gama de usuários e as mais diversas finalidades.

→ Processável por máquinas

Os dados devem ser razoavelmente estruturados de modo a permitir o processamento automatizado.

- → Não-discriminatório
 Os dados devem estar disponíveis para
 qualquer pessoa, sem necessidade de registro.
- → Não-proprietário
 Os dados devem estar disponíveis em um
 formato sobre o qual nenhuma entidade tem o controleexclusivo

→ Licença livre

Os dados não estão sujeitos a quaisquer direitos de autor, patentes, marcas comerciais ou regulamento secreto.

Pode ser permitida uma razoável privacidade e restrições de privilégio e segurança.

Portal Brasileiro de Dados Abertos

- > Portal da Transparência
- > GDELT: Os dados globais sobre eventos, localização
- > CrunchBase
- ➤ O Banco Mundial http://www.worldbank.org/
- Gapminder http://www.gapminder.org/data /
- Nações Unidas conjuntos de dados http://data.un.org/
- > Fundo Monetário Internacional http://www.imf.org/external/data.htm

→ Mineração de dados



- → UC Irvine Máquina Aprender lista de conjuntos de mineração de dados
- → KD Nuggets lista de conjuntos de mineração de dados
- → BigML lista de conjuntos de mineração de dados

- Azure Dados de mercado http://datamarket.azure.com/browse/Data
- Amazon Web Services dados públicos http://aws.amazon.com/datasets
- Google Public Data https://www.google.com/publicdata/directory
- Google Trends
 http://www.google.com/trends/explore
- Github Público conjuntos de dados https://github.com/caesar0301/awesome-
- public-datasets
- API ESPN Sports http://espn.go.com/apis/devcenter/
- Esportes Dados de Referência http://www.sports-reference.com/

Titanic

- Análise de dados do passageiros que viajavamno transatlântico Titanic
- A embarcação partiu em sua viagem inaugural de Southampton para
 - Nova lorque em 10 de abril de 1912, no caminho passando em Cherbourg-Octeville na França e por Queenstown na Irlanda.
- Ele colidiu com um iceberg às 23h40min do dia 14 de abril e afundou na madrugada do dia seguinte com mais de 1 500 pessoas a bordo, sendo um dos maiores desastres marítimos em tempos de paz de toda a história.

Ocorrências Aeronáuticas

- A base de dados de ocorrências aeronáuticas é gerenciada pelo Centro de Investigação e Prevenção de Acidentes aeronáuticos (CENIPA). Constam nesta base de dados todos os acidentes e incidentes graves notificados ao CENIPA nos últimos 10 anos e que ocorreram em solo brasileiro.
- Dentre as informações disponíveis estão os dados sobre as aeronaves envolvidas, fatalidades, localidade, data / horário e classificação / tipificação das ocorrências. São resguardados os dados pessoais de pessoas físicas / jurídicas envolvidas.

Sindec - Procon

O Sistema Nacional de Informações de Defesa do Consumidor – Sindec é um sistema informatizado que integra processos e procedimentos relativos ao atendimento de consumidores nos Procons, constituindo-se em um instrumento que permite amplificar a voz de milhões de consumidores em todo o Brasil

Comunidade no Brasil

- → 1990 usuários cadastrados (lista yahoo)
- → O grupo foi criado para geração de documentação e recursos em português do Brasil, troca de conhecimentos e experiências

https://br.groups.yahoo.com/neo/groups/pentahobr/



