



INSTITUTO POLITECNICO CASTELO BRANCO

ESCOLA SUPERIOR DE TECNOLOGIA

DISCIPLINA DE AMBIENTE INTERACTIVO

Curso de Mestrado

DESENVOLVIMENTO DE SOFTWARE E SISTEMA INTERACTIVO

Disciplina: Datawarehouse

Datawarehouse para Escola Superior de Tecnologia de Castelo Branco

Aluno:

Leandro Tito Manjate

Professora:

Filipe Fidalgo

Castelo Branco, ao fevereiro de 2022

Índice

Introdução	3
1 Definição Balance Score Cards	3
2 Modelação Multidimensional	6
2.1 Definição das SK	7
3 Processos de ETL	7
4 Resultados obtido no Excel	8
5 Geração de Relatórios	10
6 Conclusão	13
7 Bibliografia	14

Introdução

No presente relatório será feito uma investigação e análise da Escola Superior de Tecnologia de Castelo Branco com o objetivo de projetar uma datawarehouse para o mesmo poder auxiliar nas melhores tomadas de decisões possíveis para otimizar e monitorizar o desempenho da escola e assim consequentemente tentar manter os alunos internos e atrair novos alunos.

Ao longo dos anos tem se notado que os cursos de Engenharias carecem de alunos tendo em conta que maiores dos cursos tem alta empregabilidade. Normalmente as pessoas que mais frequentam os cursos são alunos estrangeiros que aparentemente aparecem no mesmo sem saber no que vão se meter e/ou provavelmente acreditam que o curso é extremamente difícil e longo desistem. Alguns alunos reclamam que não tem apoio do professor e outros as vezes não conseguem aparecer as aulas porque estão a trabalhar e as vezes quando tem tempo não aparecem porque acreditam estaria a perder o seu tempo e seria melhor estudar em casa e ir tirar duvida com o professor. Por essa razão seria criada a datawarehouse para verificar se conseguimos minimizar os problemas.

Inicialmente será falado das características identificadoras da escola que são as respetivas área de actividade, modelo de ensino, quantidade de funcionário aluno e docente, sistema de informação e o modelo de ensino. Depois será feito uma análise SWOT da escola e depois o seu respetivo alinhamento. Em seguida será feito o Balance Score cards e depois será analisada a dimensão de Sistema de Informação. Depois será feito a Modelação Multidimensional e por fim será analisado o ROI do projeto.

1 Definição Balance Score Cards

Aluno			
Objetivo	Medida	Alvo	Iniciativa
-Diminuir o número de desistência dos alunos	-Registos de dados pessoais dos alunos -Registo de motivos de desistência -Registo das frequências das aulas do aluno	-Alunos com negativas e notas baixas -Alunos com número de faltas elevada	-Reunir-se com os alunos e analisar-se o que se passa -Criação de vídeos aulas para auxílio as aulas -Disponibilização de alguma hora para ajuda os alunos
Aumentar número de estudante internacional	-Registo de número de estudante internacional atual	-Aluno de países de expressão português -Alunos de Erasmus	-Fazer anúncio nos respetivos pais -Comunicar as embaixadas de cada

	-Registo de pais que desejam captura os alunos		país anunciados vagas -
Aumentar o número de estudante de sexo feminino	-Registo de número de alunas -Registo de número total de vagas nos respetivos cursos -Registo de número de professoras	-Estudante de sexo feminino do secundário. -Mulher de sexo feminino maior de 23 anos	-Fazer palestras para tentar incentivar o publico feminino -Criar mais eventos e workshop direcionada ao publico feminino

Tabela 1 Balance Score Cards para aluno

Operativo			
Objetivo	Medida	Alvo	Iniciativa
-Criar mais curso com certificação	-Registo de procura no mercado dos respetivos cursos -Registo de empresa que faria parceria	-Alunos interno, externos e não estudantes	-Enriquecer o currículo dos alunos -Desenvolver certas competências uteis para o mercado de trabalho
-Melhorar o ranking da escola	-Registo a nível de centro -Registo a nível nacional	-Posicionar-se no top 5 melhores da escola nacional	-Aumentar o número de aluno internacional -Incentivar os alunos a criar e fazer projeto inovadores

Tabela 2 Balance Score Cards para operações

Docentes			
Objetivo	Medida	Alvo	Iniciativa
-Criar conteúdos atualizados e mais detalhados	-Registo de bibliografia atualizado -Registo de tópicos ou software mais utilizado no mercado	-Professores -Alunos interno	-Incentivar os professores a fazer os acetados e os slides
-Criar vídeos aulas	-Registo de dados de alunos da associação -Registo de horário dos alunos para filmar -Registo de modelo de camara que a escola tem disponível -Registo de aluno interno que sabem usar uma camera -Registo de horário e disponibilidade do professor	-Alunos que não aparecem em certas aulas. -Alunos que chegaram atrasado nas aulas	-Convidar alunos da associação ou alunos que não fazem parte de nenhuma associação para fazer as filmagens. -Fazer vídeos de todas aulas -O professor podia fazer um vídeo explicando cada matéria fora da aula

Tabela 3 Balance Score Cards para Docentes

Inovação			
Objetivo	Medida	Alvo	Iniciativa
-Formar os docentes/pais na utilização da plataforma	-Registo do número de participantes nas formações planeadas	-Professores	-Melhorar o processo de interação com a plataforma

--	--	--	--

Tabela 4 Balance Score Cards para inovação

2 Modelação Multidimensional

Para fazer o datawarehouse precisamos de criar uma tabela de facto e nessa mesma tabela precisamos de ter dimensões e cada dimensão tem os seus respetivos atributos. O tipo de tabela utilizado nesse caso foi uma tabela de facto sem facto que é utilizado quando não é encontrado um fato como por exemplo de venda e outras.

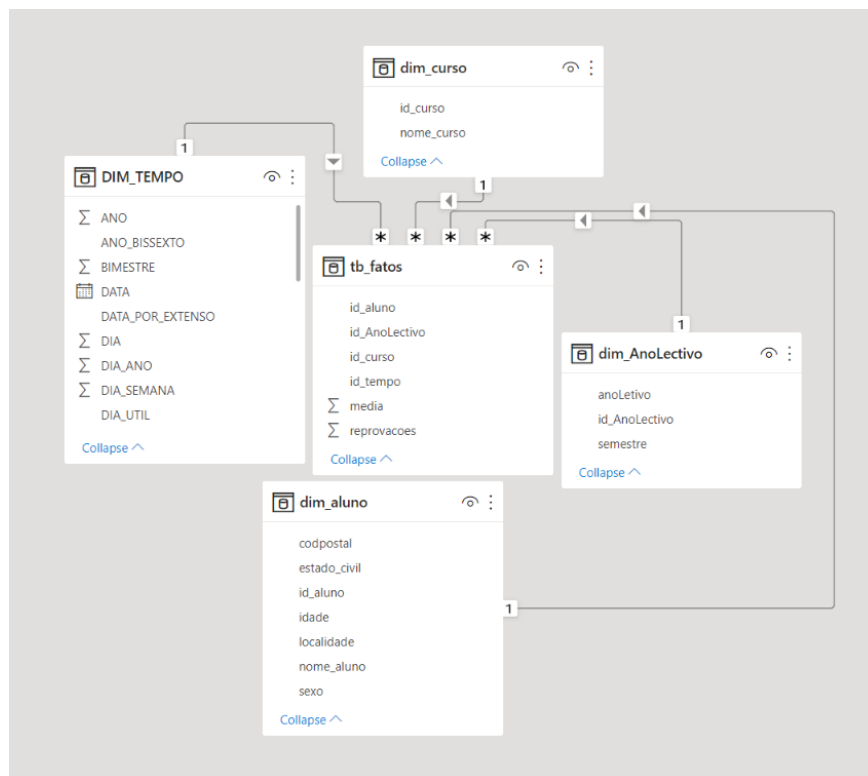


Figura 1 Tabela de facto sem facto

Com esta tabela de facto será esperado as respostas para seguintes perguntas

- Qual é numero de aluno por sexo que estão no primeiro ano que fizeram menos de 6 disciplina?
- Qual é o numero de aluno aprovados por cada ano?
- Qual é o numero de alunos com media 9.5 por ano letivo?
- Qual é o numero de alunos que concluíram o curso por curso
- Qual é o número de reprovações por ano em cada curso?

- Qual é o número de reprovações de curso por cada ano lectivo?

2.1 Definição das SK

Na nossa tabela de fatos os (surrogate keys) são os seguintes:

- **Id_aluno** que esta associado a chave primaria da tabela dimensão de aluno.
- **Id_curso** que esta associado a chave primaria da tabela dimensão de curso.
- **Id_tempo** que esta associado a chave primaria da tabela dimensão de tempo.
- **id_AnoLectivo** que esta associado a chave primaria da tabela de dimensão ano lectivo

3 Processos de ETL

Na criação de datawarehouse foi retirado informações de várias fontes como a ficha de alunos e ficha de curso e foi inserido manualmente na base de dados. Depois na base de dados transformamos as informações extraídas e dados uteis e depois a informações foi carregada para uma plataforma de *Business intelligence*.

Com a ficha de aluno utilizamos as informação para criar a tabela de dim_aluno onde no mesmo foi inserido os seguintes dados: Nome, idade, localidade, código postal, estado civil e sexo.

Com a ficha de cursos utilizamos a informação para criar a tabela de dim_curso onde no mesmo foi inserido o único dado que era o nome do curso.

Em seguida foi criado a dim_tempo que começa desde o dia 1 de janeiro de 2018 até 31 de Dezembro de 2018.

Na tabela de factos utiliza todas essas informações obtidas pelas dimensões afim de criar informações uteis para serem carregadas para um Plataforma de Visualização Power BI e o excel.

A transformação de dados foi utilizado o script que encontra no ficheiro resolucao.sql onde foi utilizado certas instruções com objetivo de obter informações uteis para responder as perguntas sugeridas.

E também certos tratamentos de dados foram feitas no excel de forma a transformar os dados em informações uteis para resolver o nosso problema



Figura 2 Processo de ETL

4 Resultados obtido no Excel

Como se sabe os alunos podem ser reprovados de formas diferente e isso pode ser por falta de presença, por não ter a media suficiente para transitar, desistência e outras. Assim foi adquirido os números de reprovações registadas em cada mês porque dependendo das escolas os registos de aprovações podem ser feitas por mês em fim de cada semestre.

Comentado [LM1]:

Comentado [LM2R1]:

Soma de reprovacoes	Rótulos de Coluna											
Rótulos de Linha	1	2	3	4	5	6	7	8	9	10	11	12 Total Geral
2018	2154	2090	2223	2261	1964	2065	2092	2757	2042	2183	1592	25825
2019	1780	1390	2149	2384	1954	1741	1682	2266	1830	2323	2125	23640
2020	2620	1998	1974	2158	2186	2008	1653	2024	1896	2778	1844	25467
2021	1981	1733	2325	1568	1427	1847	1971	2132	2099	1977	1779	22953
2022	1746	1846	2020	1788	1793	1928	1919	2122	2548	2491	1989	23992
2023	36											36
Total Geral	10317	9057	10691	10159	9324	9589	9317	11301	10415	11752	9329	121913

Figura 3 Reprovações por ano em cada mês

Na tabela os meses estão representado em numeração e pode ser feito analise de quantidades de reprovações por mês registado com o objetivo de tentar identificar a origem dos problemas.

Em 2018 podemos verificar que em agosto foi o mês onde ocorreu mais reprovações com o valor de 2757, em 2019 foi em abril com 2384 em 2020 foi em janeiro com 2620, em 2021 foi em agosto e por fim em 2022 foi em setembro.

Também podemos verificar o ano que houve mais reprovações e nesse caso foi em 2022.

Soma de reprovacoes	Rótulos de Coluna							
Rótulos de Linha	2018	2019	2020	2021	2022	2023	Total	Geral
Automação e Gestão Industrial	2135	1608	1690	1404	1719		8556	
Comunicações Móveis	1531	1738	1524	1431	1368		7592	
Construção Civil	1702	1422	1708	1597	1616		8045	
Construção Sustentável	1721	1208	1949	1500	1448		7826	
Desenvolvimento de Produtos Multimédia	1741	1539	1426	1421	1488		7615	
Desenvolvimento de Software	1129	1669	1547	1902	1403		7650	
Engenharia das Energias Renováveis	1892	1554	1613	1415	1711		8185	
Engenharia Eletrotécnica e Telecomunicações	1860	1480	1661	1419	1732		8152	
Engenharia Industrial	1688	1773	1571	1824	1580	36	8472	
Engenharia Informática	2213	1773	1898	1842	1524		9250	
Engenharia Informática	1615	1765	1878	1312	1750		8320	
Instalações Elétricas e Telecomunicações	1456	1448	1816	1473	1473		7666	
Redes e Sistemas Informáticos	1872	1426	1895	1363	1558		8114	
Tecnologias da Informação e Multimédia	1453	1582	1594	1571	1692		7892	
Tecnologias e Programação	1817	1655	1697	1479	1930		8578	
Total Geral	25825	23640	25467	22953	23992	36	121913	

Figura 4 Número de reprovação por ano em cada curso

Nessa tabela é feita analise anual de cada curso com objetivo de verificar o curso com maior número de reprovação por ano e o total de reprovações desde 2018 até 2022.

Com a tabela pode se concluir que o curso de Tecnologias e Programação possui mais reprovações.

Soma de reprovacoes	Rótulos de Coluna				
Rótulos de Linha	1	2	3	4	Total Geral
Automação e Gestão Industrial	2360	1993	2269	1934	8556
Comunicações Móveis	2023	1779	1720	2070	7592
Construção Civil	2022	1758	1882	2383	8045
Construção Sustentável	1717	2133	1979	1997	7826
Desenvolvimento de Produtos Multimédia	1820	2056	1921	1818	7615
Desenvolvimento de Software	1383	1881	2035	2351	7650
Engenharia das Energias Renováveis	1824	2058	2449	1854	8185
Engenharia Eletrotécnica e Telecomunicações	2392	1664	2094	2002	8152
Engenharia Industrial	2555	1892	2180	1845	8472
Engenharia Informática	2089	2366	2278	2517	9250
Engenharia Informática	2371	1834	2009	2106	8320
Instalações Elétricas e Telecomunicações	1928	1924	1959	1855	7666
Redes e Sistemas Informáticos	2014	2153	2079	1868	8114
Tecnologias da Informação e Multimédia	1909	1746	2365	1872	7892
Tecnologias e Programação	2457	2013	2157	1951	8578
Total Geral	30864	29250	31376	30423	121913

Figura 5 Número de reprovações por ano letivo em cada curso

Nessas tabelas foram analisados os números de reprovações por ano letivo por curso e total das reprovações em cada curso. Pode se verificar que o curso de Tecnologias e Programação era o curso com maior número de reprovações desde 2018 até 2022.

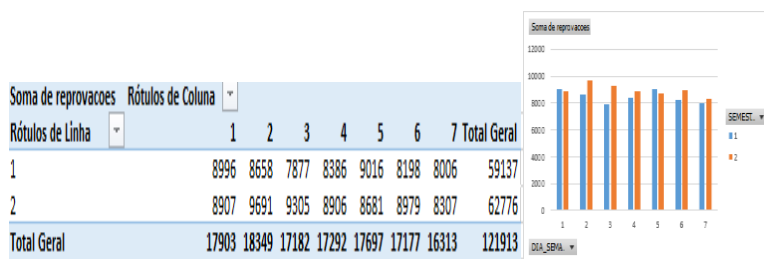


Figura 6 Número de reprovações por dia de semana por semestre

No visualizações acima pode se ver o números de reprovações registados por dia de semana por cada semestre e também o gráfico de barras para ajudar na visualização do mesmo.

5 Geração de Relatórios

Apos o carregamento das informações transformada na base de dados conseguimos gerar os seguintes dados:

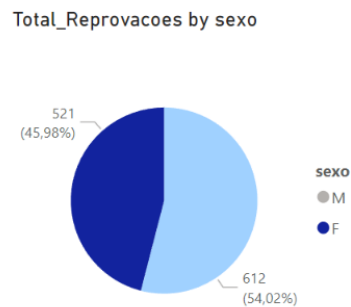


Figura 7 Numero de aluno por sexo que estão no primeiro ano que fazem menos de 6 disciplina

Com o diagrama podemos verificar que existe mais homens que não conseguem fazer a metade da disciplina no primeiro ano lectivo isso indicado que esses mesmos são os principais candidatos a desistir do curso logo no primeiro ano.

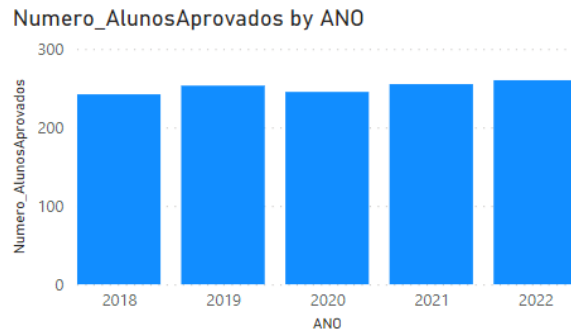


Figura 8 Numero de aluno por sexo que estão no primeiro ano que fazem menos de 6 disciplina

Pode se ver que entre 2018 até 2022 houve um ripple na taxa de aprovação mais no final a taxa de aprovação aumentou, assim indicado que algo foi bem feito nesse ano e deve ser repetitivo e melhorado.

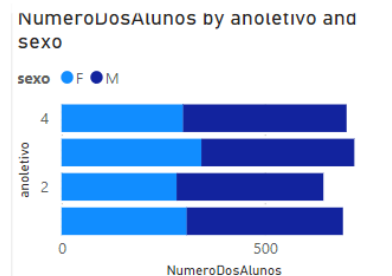


Figure 4 Numero de alunos com media 9.5 por ano lectivo

No gráfico podemos verificar os números de alunos por sexo com media menor que 9.5 em cada ano lectivo.

Com o gráfico podemos verificar que no segundo ano o sexo feminina é o sexo dominante com media menor que 9.5. No quarto ano o sexo masculino é o sexo predominante com medias menor que 9.5

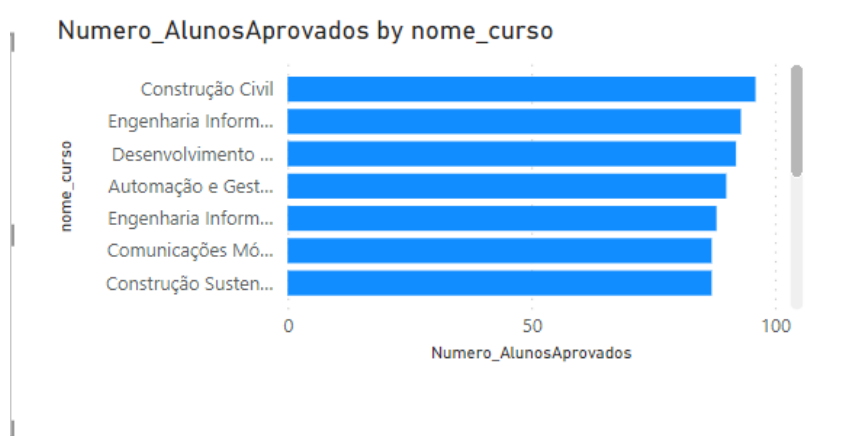


Figure 5 Numero de aprovados por cursos

Analisado o grafico podemos verificar que na construção civil existe maior taxa de aprovação do que o resto dos cursos. O curso de comunicação moveis é o curso com menos aprovações.

6 Conclusão

Com o estudo foi notado que ao criarmos o datawarehouse ajudaria muita nas decisões da escola e controlaria muito os alunos e os docentes. Tendo o software dentro de cinco anos seria possível adquirir um retorno de 72,90040889% ou um retorno maior. Com o sistema seria possível verificar se os alunos estavam satisfeitos com um determinado professor, podia se fazer um estudo se era possível criar curso com certificações caso tivesse professores qualificados e suficiente. Com o sistema conseguiríamos além de manter os alunos poderíamos aumentar o números de alunos em cada ano além disso, os alunos sairiam mais preparados porque normalmente os alunos saem da escola com falta de certos conhecimentos e depois toma uma choque quando chega no mercado de trabalho. Com os lucros seria possível investir nos Laboratório assim sendo mais fácil para os investigadores produzirem projetos de qualidade com o objetivo de dar mais reconhecimento a Escola.

7 Bibliografia

BALANCED SCORECARD INSTITUTE. (s.d.). Obtido de <https://balancedscorecard.org/bsc-basics-overview/>

Business NEWS DAILY. (s.d.). Obtido de <https://www.businessnewsdaily.com/4245-swot-analysis.html>

Diário de Notícias. (s.d.). Obtido de <https://www.dn.pt/portugal/portugal-gasta-6200-eurosano-por-aluno-no-publico-4114874.html>

EDULOG. (s.d.). Obtido de <https://www.edulog.pt/artigos/em-analise/os-paises-da-ocde-gastam-em-media-9184-eurosano-por-aluno>

Expresso. (s.d.). Obtido de <https://expresso.pt/sociedade/2021-09-13-Cada-aluno-na-escola-publica-custa-6200-euros-por-ano-ha-seis-anos-custava-4700-E-um-aumento-brutal-diz-ministro-da-Educacao-0a8b242d>

Instituto Politécnico Castelo Branco. (s.d.). Obtido de Corpo de Gestão: <https://www.ipcb.pt/estcb/org%C3%A3os-de-gest%C3%A3o>

Kimbal, R., Reeves, L., Ross, M., & Thornthwaite, W. (s.d.). *Datawarehouse Lifecycle Toolkit.*

LivePlan. (s.d.). Obtido de <https://www.liveplan.com/blog/what-is-a-swot-analysis-and-how-to-do-it-right-with-examples/>

NOS. (s.d.). Obtido de <https://www.nos.pt/empresas/repositorio-informacao/criar-uma-empresa/guias-teoricos/Pages/balanced-scorecard.aspx>

Público. (s.d.). Obtido de <https://www.publico.pt/2021/09/17/sociedade/noticia/custo-aluno-dividese-orcamento-total-estudantes-1977874>

RESULTADOS DIGITAIS. (s.d.). Obtido de <https://resultadosdigitais.com.br/blog/o-que-e-roi-retorno-sobre-investimento/>

Sapo. (s.d.). Obtido de <https://poligrafo.sapo.pt/fact-check/custo-medio-por-aluno-no-ensino-publico-e-superior-as-propinas-nos-melhores-colegios-privados>

Universidade de Coimbra. (s.d.). Obtido de https://estudogeral.uc.pt/bitstream/10316/24772/1/Tatiana_Ferreira_Relat%c3%b3rio%20de%20Est%c3%a1gio_FEUC-2013.pdf

WordStream. (s.d.). Obtido de <https://www.wordstream.com/blog/ws/2017/12/20/swot-analysis>